

ISSN 1684-8853 (print); ISSN 2541-8610 (online)

ИНФОРМАЦИОННО- УПРАВЛЯЮЩИЕ СИСТЕМЫ

НАУЧНЫЙ ЖУРНАЛ

5(120)/2022

5(120)/2022

PEER REVIEWED JOURNAL

INFORMATSIONNO- UPRAVLIAYUSHCHIE SISTEMY (INFORMATION AND CONTROL SYSTEMS)

Founder

A. Vostrikov

PublisherSaint Petersburg State University
of Aerospace Instrumentation**Editor-in-Chief**

E. Krouk

Dr. Sc., Professor, Moscow, Russia

Executive secretary

O. Muravtsova

Editorial Board

S. Andreev

Dr. Sc., Tampere, Finland

V. Anisimov

Dr. Sc., Professor, Saint Petersburg, Russia

B. Bezruchko

Dr. Sc., Professor, Saratov, Russia

N. Blaunstein

Dr. Sc., Professor, Beer-Sheva, Israel

M. Buzdalov,

PhD, Researcher, Saint Petersburg, Russia

C. Christodoulou

PhD, Professor, Albuquerque, New Mexico, USA

A. Dudin

Dr. Sc., Professor, Minsk, Belarus

I. Dumer

PhD, Professor, Riverside, USA

M. Favorskaya

Dr. Sc., Professor, Krasnoyarsk, Russia

L. Fortuna

PhD, Professor, Catania, Italy

A. Fradkov

Dr. Sc., Professor, Saint Petersburg, Russia

A. Hramov

Dr. Sc., Professor, Innopolis, Russia

L. Jain

PhD, Professor, Canberra, Australia

G. Matvienko

Dr. Sc., Professor, Tomsk, Russia

A. Myllari

PhD, Professor, Grenada, West Indies

K. Samouylov

Dr. Sc., Professor, Moscow, Russia

J. Seberry

PhD, Professor, Wollongong, Australia

M. Sergeev

Dr. Sc., Professor, Saint Petersburg, Russia

A. Shalyto

Dr. Sc., Professor, Saint Petersburg, Russia

A. Shepeta

Dr. Sc., Professor, Saint Petersburg, Russia

Yu. Shokin

RAS Academician, Dr. Sc., Novosibirsk, Russia

A. Smirnov

Dr. Sc., Professor, Saint Petersburg, Russia

T. Sutikno

PhD, Associate Professor, Yogyakarta, Indonesia

Z. Yuldashev

Dr. Sc., Professor, Saint Petersburg, Russia

R. Yusupov

RAS Corr. Member, Dr. Sc., Professor, Saint Petersburg, Russia

A. Zeifman

Dr. Sc., Professor, Vologda, Russia

Editor: A. Larionova**Proofreader:** T. Zvertanovskaia**Design:** M. Chernenko, Yu. Umnitsyna**Layout and composition:** Yu. Umnitsyna**Contact information**The Editorial and Publishing Center, SUAI
67A, Bol'shaya Morskaya, 190000, Saint Petersburg, RussiaWebsite: <http://i-us.ru/en>, e-mail: i-us.spb@gmail.com

Tel.: +7 - 812 494 70 02

INFORMATION PROCESSING AND CONTROL*Tyryshkina Y. S., Tumkovskiy S. R. A method to accelerate the joining of distributed datasets by a given criterion*

2

INFORMATION AND CONTROL SYSTEMS*Dvoynikova A. A., Kagirov I. A., Karpov A. A. Analytical review of methods for automatic detection of user engagement in virtual communication*

12

Isaeva O. S., Isaev S. V., Kulyasov N. V. Formation of adaptive publications from the Internet of Things data broker

23

HARDWARE AND SOFTWARE RESOURCES*Aderkina A. A., Sinitsyn A. S. Development of an automated software tool based on machine learning methods for solving problems of radio planning in subway sections*

32

INFORMATION SECURITY*Pimenov A. V., Fedorov I. R., Bezzateev S. V. Designing fog computing architecture with the use of blockchain technology*

40

INFORMATION CODING AND TRANSMISSION*Burkov A. A., Rachugin R. O., Turlikov A. M. Analyzing and stabilizing multichannel ALOHA with the use of the preamble-based exploration phase*

49

INFORMATION ABOUT THE AUTHORS

60

5(120)/2022

ИНФОРМАЦИОННО-
УПРАВЛЯЮЩИЕ
СИСТЕМЫ

РЕЦЕНЗИРУЕМОЕ ИЗДАНИЕ

Учредитель

А. А. Востриков

ИздательСанкт-Петербургский государственный университет
аэрокосмического приборостроения**Главный редактор**Е. А. Крук,
д-р техн. наук, проф., Москва, РФ**Ответственный секретарь**

О. В. Муравцова

Редакционная коллегия:

С. Д. Андреев,

д-р техн. наук, Тампере, Финляндия

В. Г. Анисимов,

д-р техн. наук, проф., Санкт-Петербург, РФ

Б. П. Безручко,

д-р физ.-мат. наук, проф., Саратов, РФ

Н. Блаунштейн,

д-р физ.-мат. наук, проф., Беэр-Шева, Израиль

М. В. Буэдалов,

канд. техн. наук, научный сотрудник, Санкт-Петербург, РФ

Л. С. Джайн,

д-р наук, проф., Канберра, Австралия

А. Н. Дудин,

д-р физ.-мат. наук, проф., Минск, Беларусь

И. И. Думер,

д-р наук, проф., Риверсайд, США

А. И. Зейфман,

д-р физ.-мат. наук, проф., Вологда, РФ

К. Кристоделу,

д-р наук, проф., Альбукерке, Нью-Мексико, США

Г. Г. Матвиенко,

д-р физ.-мат. наук, проф., Томск, РФ

А. А. Мюллери,

д-р наук, профессор, Гренада, Вест-Индия

К. Е. Самуйлов,

д-р техн. наук, проф., Москва, РФ

Д. Себерри,

д-р наук, проф., Волонгонг, Австралия

М. Б. Сергеев,

д-р техн. наук, проф., Санкт-Петербург, РФ

А. В. Смирнов,

д-р техн. наук, проф., Санкт-Петербург, РФ

Т. Сутикнуо,

д-р наук, доцент, Джокьякарта, Индонезия

М. Н. Фаворская,

д-р техн. наук, проф., Красноярск, РФ

Л. Фортуна,

д-р наук, проф., Катания, Италия

А. Л. Фрадков,

д-р техн. наук, проф., Санкт-Петербург, РФ

А. Е. Храмов,

д-р физ.-мат. наук, Иннополис, РФ

А. А. Шалыто,

д-р техн. наук, проф., Санкт-Петербург, РФ

А. П. Шепета,

д-р техн. наук, проф., Санкт-Петербург, РФ

Ю. И. Шокин,

акад. РАН, д-р физ.-мат. наук, проф., Новосибирск, РФ

З. М. Юлдашев,

д-р техн. наук, проф., Санкт-Петербург, РФ

Р. М. Юсупов,

чл.-корр. РАН, д-р техн. наук, проф., Санкт-Петербург, РФ

Редактор: А. Г. Ларионова**Корректор:** Т. В. Звертановская**Дизайн:** М. Л. Черненко, Ю. В. Умницына**Компьютерная верстка:** Ю. В. Умницына**Адрес редакции:** 190000, г. Санкт-Петербург,

ул. Большая Морская, д. 67, лит. А, ГУАП, РИЦ

Тел.: (812) 494-70-02, эл. адрес: ius.spb@gmail.com,

сайт: <http://i-us.ru>**ОБРАБОТКА ИНФОРМАЦИИ И УПРАВЛЕНИЕ***Тырышкина Е. С., Тумковский С. Р. Метод ускорения объединения
распределенных наборов данных по заданному критерию*

2

ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ*Двойникова А. А., Кагиров И. А., Карпов А. А. Аналитический обзор
методов автоматического распознавания вовлеченности
пользователя в виртуальную коммуникацию*

12

*Исаева О. С., Исаев С. В., Кулясов Н. В. Формирование адаптивных
рассылок брокера данных интернета вещей*

23

ПРОГРАММНЫЕ И АППАРАТНЫЕ СРЕДСТВА*Адёркина А. А., Синицын А. С. Разработка автоматизированного
программного инструмента на основе методов машинного
обучения для решения задач радиопланирования
на участках метрополитена*

32

ЗАЩИТА ИНФОРМАЦИИ*Пименов А. В., Федоров И. Р., Беззатеев С. В. Построение
архитектуры туманных вычислений с использованием технологии
блокчейн*

40

КОДИРОВАНИЕ И ПЕРЕДАЧА ИНФОРМАЦИИ*Burkov A. A., Rachugin R. O., Turlikov A. M. Analyzing and stabilizing
multichannel ALOHA with the use of the preamble-based
exploration phase*

49

СВЕДЕНИЯ ОБ АВТОРАХ

60

Журнал входит в БД Scopus и в Перечень рецензируемых научных изданий,
в которых должны быть опубликованы основные научные результаты диссертаций
на соискание ученой степени кандидата наук,
на соискание ученой степени доктора наук.

Сдано в набор 05.09.22. Подписано в печать 27.10.22. Дата выхода в свет: 31.10.2022.

Формат 60×841/8. Гарнитура CentSchbkCyril BT. Печать цифровая.

Усл. печ. л. 7,4. Уч.-изд. л. 10,2. Тираж 1000 экз (1-й завод 50 экз.). Заказ № 541.

Оригинал-макет изготовлен в редакционно-издательском центре ГУАП.

190000, г. Санкт-Петербург, ул. Большая Морская, д. 67, лит. А.

Отпечатано в редакционно-издательском центре ГУАП.

190000, г. Санкт-Петербург, ул. Большая Морская, д. 67, лит. А.

Распространяется бесплатно.

Журнал зарегистрирован в Министерстве РФ по делам печати,
телерадиовещания и средств массовых коммуникаций.
Свидетельство о регистрации ПИ № 77-12412 от 19 апреля 2002 г.
Перерегистрирован в Роскомнадзоре.
Свидетельство о регистрации ПИ № ФС77-82226 от 23 ноября 2021 г.

© А. А. Востриков, 2022



Метод ускорения объединения распределенных наборов данных по заданному критерию

Е. С. Тырышкина^а, аспирант, orcid.org/0000-0003-4814-4874

С. Р. Тумковский^а, доктор техн. наук, доцент, orcid.org/0000-0002-1647-2171, STumkovskiy@hse.ru

^аНациональный исследовательский университет «Высшая школа экономики», Мясницкая ул., 20, Москва, 101000, РФ

Введение: быстро растущие объемы информации бросают новые вызовы современным технологиям анализа данных. Одной из самых распространенных сопутствующих операций в аналитике является объединение наборов данных. Объединение — крайне ресурсоемкая операция, которая тяжело поддается масштабируемости и повышению эффективности использования ресурсов в распределенных базах данных или системах, основанных на парадигме MapReduce. **Цель:** разработать метод, позволяющий ускорить объединение наборов данных в распределенных системах. **Результаты:** рассмотрены архитектура Apache Spark, особенности распределенных вычислений на основе MapReduce, проанализированы типовые методы объединения наборов данных, рассмотрены основные способы ускорения операции объединения данных. Предложен метод, позволяющий ускорить частный случай объединения, реализованный в Apache Spark с использованием приемов партиционирования и частичной передачи наборов на вычислительные узлы кластера, таким образом, чтобы задействовать одновременно преимущества merge и broadcast объединений. Представленные экспериментальные данные демонстрируют, что метод тем значительно ускоряет операцию объединения данных относительно стандартных методов, чем больше объем входных данных. Так, для 2 ТБ сжатых данных было получено ускорение до ~37 % в сравнении со стандартным механизмом Spark SQL.

Ключевые слова — распределенные вычисления, MapReduce, Apache Spark, объединение данных.

Для цитирования: Тырышкина Е. С., Тумковский С. Р. Метод ускорения объединения распределенных наборов данных по заданному критерию. *Информационно-управляющие системы*, 2022, № 5, с. 2–11. doi:10.31799/1684-8853-2022-5-2-11, EDN: AAGGGR

For citation: Tyryshkina Y. S., Tumkovskiy S. R. A method to accelerate the joining of distributed datasets by a given criterion. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 5, pp. 2–11 (In Russian). doi:10.31799/1684-8853-2022-5-2-11, EDN: AAGGGR

Введение

В течение многих лет подходы к оптимизации запросов заимствовались из традиционных баз данных и адаптировались к среде распределенных баз данных. Таким образом, знания, полученные при работе с централизованными реляционными системами управления базами данных, были использованы для оптимизации вычислений в распределенных системах. Однако распределенные вычисления обладают уникальными характеристиками, которые часто создают как новые проблемы, так и новые возможности оптимизации.

Одной из важных задач при анализе данных является задача их объединения. В настоящее время она широко изучается, предлагаются новые методы ее решения [1–3], особенно для работы с большими наборами данных в распределенной среде, однако имеющиеся в настоящее время методы недостаточно эффективны: требуют больших перемещений данных между вычислительными нодами кластера, используют много оперативной памяти.

Первая проблема, с которой приходится сталкиваться при выполнении запроса на объединение данных в распределенных базах данных, — это сортировка и перемещение данных внутри кластера, которое увеличивает нагрузку на сеть, что, несмотря на технический прогресс, остается одним из узких мест в кластерных системах. Вторая проблема заключается в том, что невозможно поместить весь набор данных в оперативную память, и данные многократно перезаписываются на жесткие диски, что также сильно замедляет их обработку. Третья проблема вызвана особенностями данных, а именно тем, что некоторые значения в данных встречаются гораздо чаще, чем другие. В этом случае простая стратегия объединения является причиной неравномерного разделения данных на партиции на вычислительных узлах. Затраты на обработку таких данных приводят в лучшем случае к значительному увеличению общего времени решения задачи объединения, а в худшем — к полной невозможности ее решить.

В настоящей работе рассматриваются особенности вычислений в распределенной среде, обсужда-

ется архитектура Apache Spark, для которой предложена стратегия ускорения решения задачи объединения данных. В качестве основного критерия сравнения архитектур Apache Spark и MapReduce использовано время решения задачи объединения данных. При этом надежность вычислений и удобство разработки считаются равными. Показаны возможные проблемы и узкие места типовых стратегий объединения данных и особенности их применения. Представлен разработанный метод решения задачи объединения распределенных наборов данных по заданному критерию и его экспериментальные исследования.

Анализ технологии MapReduce

В настоящее время крупные аналитические платформы ежедневно обрабатывают десятки терабайт данных. Это привело к появлению новых файловых систем и систем управления базами данных [4]. Одной из таких является файловая система Google (GFS) [5] или распределенная файловая система Hadoop (HDFS). Долгое время самым быстрым и надежным решением для создания приложений в этой файловой системе был Google MapReduce, который является основным компонентом Hadoop и обеспечивает обработку данных [6]. Hadoop работает со структурированными, слабоструктурированными и неструктурированными данными.

MapReduce — инструмент распределенных вычислений, способный работать на тысячах узлов кластера [7], является моделью программирования для разработки масштабируемых параллельных приложений с большими данными в вычислительных кластерах [8].

Концепция MapReduce заключается в том, что процесс обработки данных, хранящихся в одном или нескольких файлах, делится на два этапа: отображение и сокращение (рис. 1). Результатом обработки являются пары ключ-значение. Каждый этап имеет ключ-значение в качестве входа и выхода. Ключи и значения не являются внутренними свойствами данных, но они вы-

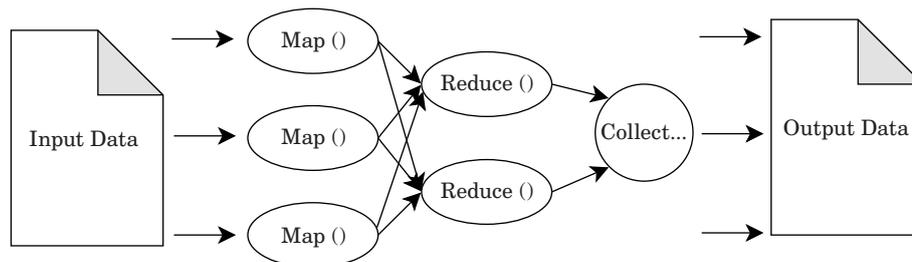
бираются пользователем, который анализирует данные. Генерация пары ключ-значение Hadoop зависит от набора входных данных и требуемого выхода. Основная задача первого этапа (map) — отфильтровать данные и предоставить возможность группировать данные по ключу, т. е. $(k1, v1) \rightarrow List(k2, v2)$. Второй этап (reduce) использует функцию для объединения значений, связанных одним ключом, и получения результата для этого ключа, т. е. $(k2, List(v2)) \rightarrow List(k3, v3)$. Ключ — это поле или объект, по которому будет выполняться группировка и агрегация на этапе reduce. Значение — это поле или объект, к которому будет применяться функция агрегирования.

Данные на выходе из этапа map разделяются и сортируются таким образом, чтобы их можно было распределить в соответствии с критерием группировки на следующей стадии по датаграммам, которые будут участвовать в вычислении агрегатной функции на этапе reduce.

Что касается решения задачи объединения данных [9], технология MapReduce гарантирует высокую надежность вычислений. Чрезмерно медленное решение задачи из-за неравномерности ключей в партиционированных наборах данных и многократного повторения операций чтения и записи является недостатком данной технологии [10, 11].

Анализ архитектуры Spark

Apache Spark — это технология кластерных вычислений, получившая развитие на основе технологии MapReduce. В отличие от заложенной в модель MapReduce концепции на основе этапов map и reduce, которая использует дисковое хранилище, в Spark применяется подход поэтапной обработки в оперативной памяти. Это позволяет существенно увеличить скорость обработки в Spark по сравнению с MapReduce [12]. Использование оперативной памяти в процессе вычислений предоставляет возможность многократного доступа к данным, что необходимо для быстрой



■ **Рис. 1.** Диаграмма MapReduce
 ■ **Fig. 1.** MapReduce diagram

работы алгоритмов машинного обучения, где широко применяются итерационные алгоритмы. Кеширование промежуточных результатов в памяти ускоряет работу таких алгоритмов.

В основе архитектуры Apache Spark лежат две основные абстракции:

- 1) направленный ациклический граф (Directed Acyclic Graphs, DAG);
- 2) устойчивые распределенные наборы данных (Resilient Distributed Datasets, RDD).

Directed Acyclic Graphs – это последовательность вычислений, выполняемых с данными, где каждая вершина является разделом RDD, а ребра – это функция преобразования данных. Здесь под разделами понимаются части данных, лежащих на разных датанодах. В отличие от партиции, данные в разделе RDD не объединяются некоторым общим критерием.

Resilient Distributed Datasets – это интерфейс для работы с коллекцией неизменяемых объектов Java или Scala, которые являются ссылками на разделы набора данных, распределенные по нескольким узлам кластера. В процессе работы разделы могут быть загружены в память на датанодах кластера, а коллекции являются постоянными для того, чтобы их можно было восстановить в случае потери части набора данных. Абстракция DAG помогает устранить многоэтапную модель выполнения Hadoop MapReduce и обеспечивает повышение производительности по сравнению с Hadoop. Весь процесс работы с данными в Spark заключается в применении к данным двух типов операций: преобразования и действия.

Apache Spark использует архитектуру главный/подчиненный с двумя основными программами-демонами (Master Daemon, Slave Daemon) и менеджером кластера (Cluster Manager). В кластере Spark всегда есть один главный (драйвер) и любое количество подчиненных (исполнителей) (рис. 2). Драйвер и исполнители запускают свои отдельные процессы Java, и пользователи

могут запускать их в одном кластере Spark или на разных машинах.

Программа-драйвер, исполняемая на главном узле кластера Spark, планирует задание и координирует его с диспетчером кластера. Драйвер преобразует RDD в DAG, разбивает задачу на несколько этапов и сохраняет метаданные обо всех RDD и их разделах.

Исполнитель отвечает за выполнение задач и обрабатывает данные. Исполнитель взаимодействует с системами хранения: читает и записывает данные во внешние источники, сохраняет результаты вычислений в памяти, кеше или на жестких дисках.

При запуске приложения Spark драйвер преобразует пользовательский код в логический DAG. На этом этапе драйвер пытается оптимизировать DAG, на основе которого строится физический план выполнения программы. Программа разбивается на небольшие задачи и отправляется на выполнение в кластер. После этого драйверу необходимо согласовать использование ресурсов с менеджером кластера, который запускает процессы исполнителя на узлах кластера.

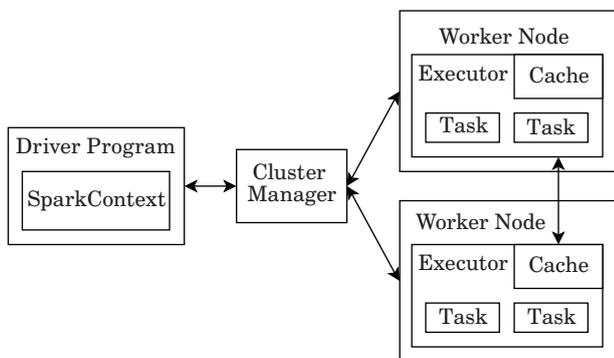
Распространенные методы выполнения операции объединения в Apache Spark

Обычно на операции объединения требуется много ресурсов кластера как сетевых, поскольку они вызывают перераспределение данных внутри кластера, так и вычислительных, чтобы провести сортировку записей и их локальное объединение, т. е. выполнить в памяти много операций сравнения [14]. Если RDD не имеют схемы партиционирования или они различны для каждого набора данных, их необходимо будет создать, что приведет к перемещению данных таким образом, чтобы оба RDD имели одинаковые схемы и данные с одинаковыми ключами находились в одних и тех же партициях (рис. 3).

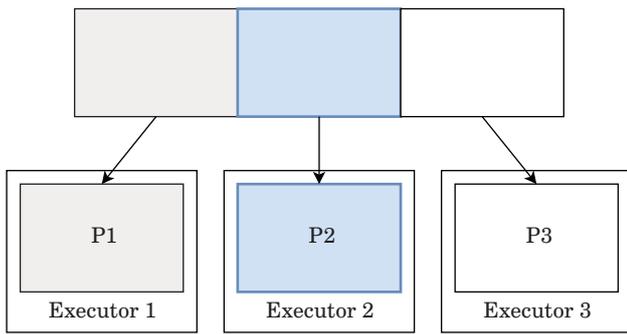
Как и в случае с большинством операций, ключ-значение, время, затраченное на выполнение операции, нагрузка на сеть между нодами кластера и количество ресурсов памяти увеличиваются с количеством ключей и расстоянием передачи данных, т. е. физической протяженностью маршрута, в том числе количеством маршрутизирующего и регенерирующего оборудования на участке сети, который записи должны преодолеть внутри кластера, чтобы добраться до нужной партиции.

Apache Spark использует в своих вычислениях две стратегии [14]: узел-узел и один узел.

Стратегия объединения узел-узел предполагает, что данные во время операции объедине-



■ **Рис. 2.** Архитектура Apache Spark [13]
 ■ **Fig. 2.** Apache Spark architecture [13]



■ **Рис. 3.** Схема идеального партиционирования, при которой данные равномерно разделены на партии и каждая партия (P) целиком находится на одном исполнителе (Executor)

■ **Fig. 3.** The scheme of ideal partitioning, in which the data is evenly divided into partitions and each partition (P) is entirely located on one executor

ния будут перемещаться между узлами кластера так, чтобы все ключи оказались на одном узле, в то время как стратегия одного узла выполняет широковещательный тип объединения, когда данные одного набора данных отправляются на каждый узел, т. е. информация дублируется на каждой датаноде в кластере, который уже может сравнивать ключи локально. Эти подходы будут рассмотрены более подробно ниже.

Способы ускорения операции объединения в Apache Spark

В первую очередь необходимо убедиться, что все основные рекомендации по выполнению операции объединения уже выполнены. Перед объединением следует подготовить данные:

– если оба RDD имеют одинаковые ключи партиционирования, то для того, чтобы ускорить операцию, желательно сначала выполнить отдельную операцию агрегирования (combByKey, cogroup);

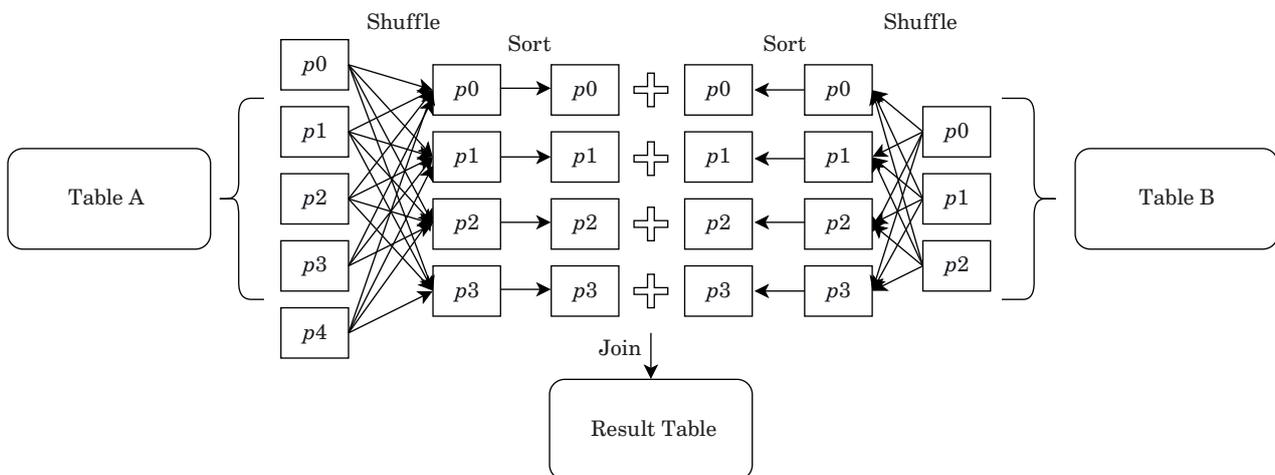
– если один RDD имеет некое легко идентифицируемое подмножество ключей, которые вы в конечном итоге не будете использовать, то лучше их отфильтровать перед объединением;

– предварительно репартиционировать данные, если это возможно.

Выделяются три основные стратегии объединения: Sort-Merge, Broadcast и Shuffle Hash.

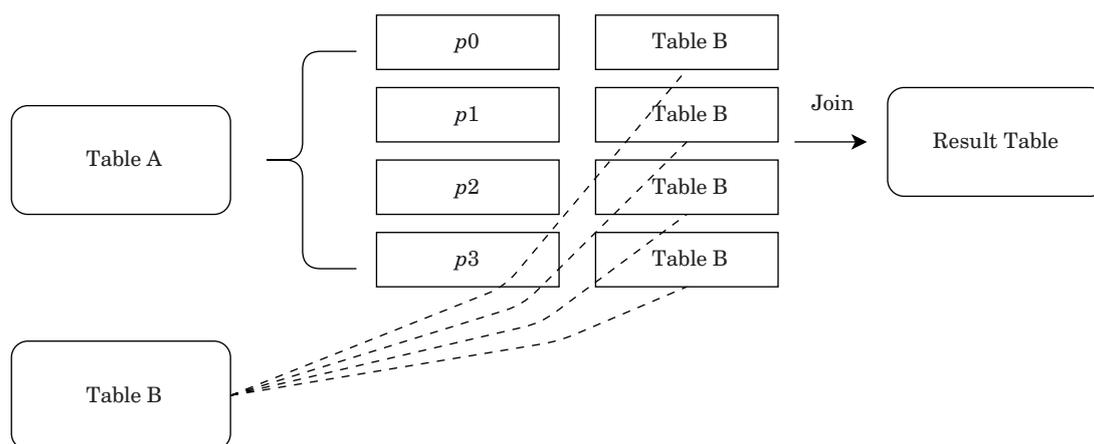
Объединение Sort-Merge (рис. 4) состоит из двух этапов и является предпочтительным, поскольку может записывать данные на диски, не используя хранение в оперативной памяти. На первом этапе сортируются наборы данных, что само по себе занимает много времени. Ускорение данного этапа исследовалось в работах [15–17]. На втором этапе отсортированные данные объединяются в партии и поэлементно сравниваются по ключу. В Spark выше версии 2.3 Sort-Merge используется в качестве алгоритма по умолчанию. Однако его использование по умолчанию можно отключить с помощью параметра конфигурации spark.sql.join.preferSortMergeJoin.

Объединение Broadcast (рис. 5) обеспечивает максимальную производительность среди стандартных методов и решает проблемы неравномерного сегментирования и ограниченного параллелизма, однако он актуален только для небольших наборов данных. Чтобы применить этот подход, размер таблицы должен быть меньше значения, настроенного с помощью конфигурации контекста spark.sql.autoBroadcastJoinThreshold (по умолчанию 10 МБ).



■ **Рис. 4.** Объединение Sort-Merge

■ **Fig. 4.** Sort-Merge join diagram



■ **Рис. 5.** Объединение Broadcast
 ■ **Fig. 5.** Broadcast join diagram

Объединение Shuffle Hash (рис. 6) работает на основе концепции MapReduce. На этапе map на основе данных вычисляются ключи, по которым будет выполняться объединение. Эти значения используются как ключи партиции, а данные перемещаются внутри кластера так, чтобы записи, относящиеся к одному ключу объединения, были расположены на одной датаноде. Spark выбирает Shuffle Hash объединение, когда невозможно применить Sort-Merge, что проверяется функцией:

```
def canBuildLocalHashMap (plan: LogicalPlan): Boolean = {
    plan.statistics.sizeInBytes < conf.autoBroadcastJoinThreshold
    * conf.numShufflePartitions
}
```

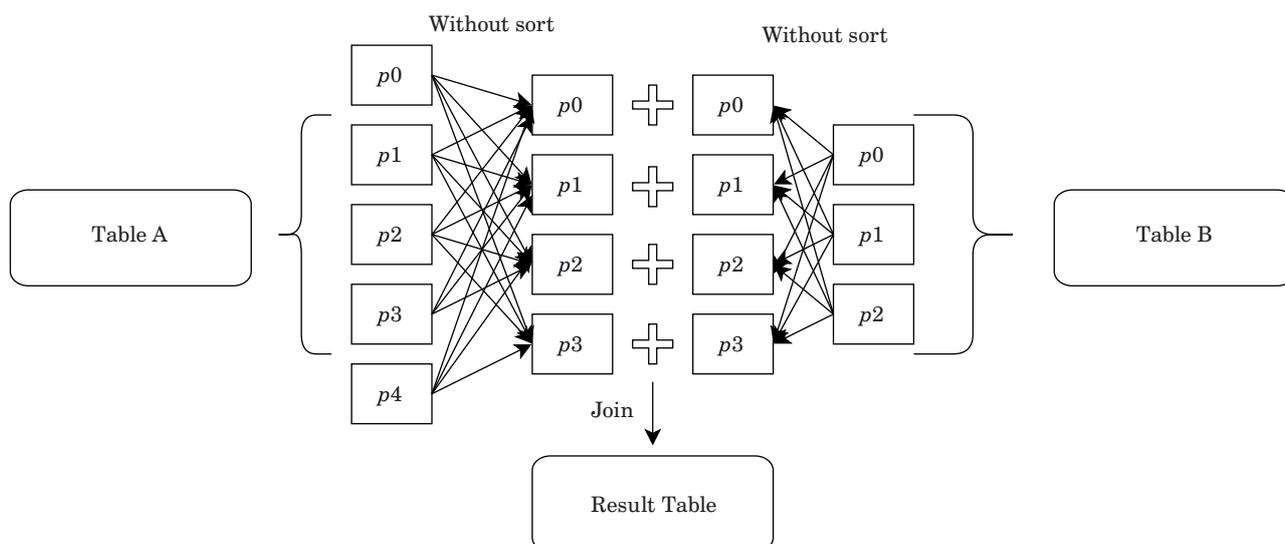
Создание хеш-таблиц — сложная операция, и ее можно выполнить, только если средний раз-

мер одной партиции достаточно мал, настолько, чтобы поместиться в оперативную память вычислительной машины, на которой производится расчет.

На практике в обычном режиме использования кластера производительность также зависит от конфигурации Spark, посторонней нагрузки на кластер в момент расчета и синергии между конфигурацией и кодом.

Метод ускорения операции объединения

Предлагаемый в работе метод сочетает в себе преимущества вышеперечисленных подходов. В то же время он не требует, чтобы наборы данных полностью помещались в ОЗУ и проводи-



■ **Рис. 6.** Объединение Shuffle Hash
 ■ **Fig. 6.** Shuffle Hash diagram

лась их сортировка, что обеспечивает простоту и более высокую скорость выполнения операции объединения. Допустим, у нас есть два несортированных набора данных — RDD1 и RDD2. Необходимо позаботиться о том, чтобы набор RDD1 хранился на диске таким образом, чтобы данные каждой партии записывались в отдельный файл. Это можно сделать, описав свой собственный класс `customPartitioner`, который задает некоторую логику партиционирования, например, вычисляя хеш-функцию по критерию объединения. Однако при слишком большом количестве партий планирование задач займет больше времени, чем фактический расчет, а при слишком маленьком будет низкая параллельность расчета. Важно равномерно распределить набор данных по партициям и предотвратить неравномерное их разделение, когда большой объем данных концентрируется на одной или нескольких машинах в кластере [18]. Допустимый для работы метода процент неравномерности данных ограничивается физическими ресурсами кластера. Данных в любой партии всегда должно быть меньше, чем имеется оперативной памяти на узлах кластера. Неравномерное попадание записей в одну партию может привести к ошибкам OOM (Out of Memory) на исполнителях. Также если данные сильно неравномерны по ключу [19], то один исполнитель может работать дольше, чем остальные, и задержать общее выполнение задачи. Таким образом, неравномерность распределения данных приводит к пропорциональному увеличению времени расчета. В данном случае проблема неравномерности решалась путем прогнозирования и распределения проблемных ключей на основе информации, предварительно полученной из данных, и наложения на них дополнительной функции партиционирования. Причина неравномерности данных всегда обусловлена семантикой, источником данных и критерием объединения. Поэтому в процессе решения задачи объединения стоит исследовать, насколько равномерно данные группируются по критерию будущего объединения. Пример того, как в Spark можно записать партии в один файл:

```
RDD.map{ line => someFunc
}.partitionBy(customPartitioner).saveAsTextFile(filePath)
```

Чтобы исполнители самостоятельно скачивали файл из HDFS, необходимо передать конфигурацию Hadoop на каждый узел, участвующий в расчетах:

```
val confBroadcast = sc.broadcast(new SerializableWritable(sc.hadoopConfiguration))
def readFromHDFS(configuration: Configuration, path: String): String = {
  val fs: FileSystem = FileSystem.get(configuration)
```

```
  val inputStream = fs.open(new Path(path));
  val writer = new StringWriter();
  IOUtils.copy(inputStream, writer, "UTF-8");
  writer.toString
}
```

Ключевым преимуществом этого подхода является отсутствие этапа `shuffle`, задача которого — промежуточное перемещение данных внутри кластера с целью сконцентрировать данные на основе ключа партии на одном узле, т. е. физическое перемещение данных между нодами кластера во время разделения данных на партии. Операция `saveAsTextFile` при RDD1 производится без этапа `shuffle`, а партиционирование RDD2 существует только на логическом уровне программы.

Таким образом:

- RDD1 делится на партии путем записи данных каждой партии в отдельный файл;
- RDD2 никак не перемещается внутри кластера, так как объекты этого RDD ссылаются на партии локально на каждой дананод. Строки, принадлежащие одной партии данного датасета, не концентрируются на одном узле, а остаются распределенными по разным узлам.

Поскольку известно, в каком файле находятся строки определенной партии RDD1, то мы можем прочитать этот файл напрямую на том узле, где находится соответствующая партия RDD2. Преимущество метода в скорости получается за счет того, что, в отличие от стандартных стратегий объединения, в нашем случае исключается промежуточный этап перемещения данных между узлами, а между нодами кластера перемещается только часть данных на объединение — RDD1, в то время как RDD2 физически не перемещается.

Для работы с партиями используются функции `mapPartitions()` или `mapPartitionsWithIndex()`, которые преобразуют каждую партию исходного набора данных в несколько элементов результата (в некоторых случаях ни в один). Одним из важных вариантов использования этих функций может быть некоторая тяжелая инициализация, которая делается один раз для нескольких элементов и повторно используется для каждого потока/партии, в отличие от функции `map()`, применяемой к каждому элементу в наборе данных.

Затем в процессе обработки RDD2 можно произвести переразбиение по тому же алгоритму, что и в RDD1 (с помощью `customPartitioner`), а затем внутри функции `mapPartitionsWithIndex()` обратиться по индексу к конкретному файлу первого набора и загрузить его в память напрямую с использованием структуры данных «словарь». Словари реализованы в виде хеш-таблиц, где операции поиска выполняются за постоянное

время, так как элемент запрашивается напрямую, без проведения сравнений, т. е. сложность алгоритма поиска соответствует $O(1)$:

```
val data = RDD2.partitionBy(customPartitioner) val res = data.
mapPartitionsWithIndex((index, partition) = {
val conf = confBroadcast.value.value
val partitionRDD1 = readFromHDFS(conf, path + index)
val partitionData = partitionRDD1.map {
case Array(k, v) => k -> v.toLong;
case _ => "" -> ""}
}.toMap
val newPartition = partition.map(
record => {(record, partitionData(record._1))})
new Partition
})
```

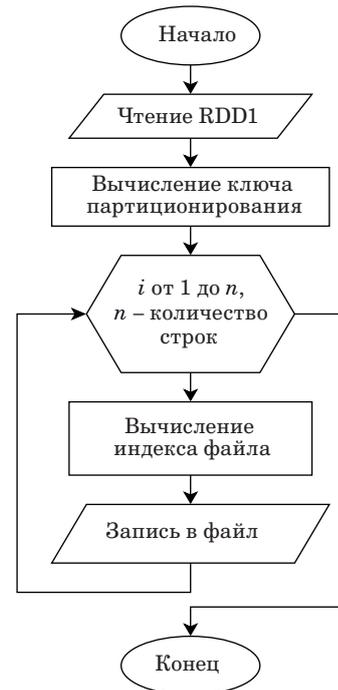
Важно помнить, что процедура переразбиения выполнятся вне контекста Spark. Spark не сможет отслеживать и предоставлять статистику загрузки данных, поэтому разработчик должен самостоятельно контролировать использование оперативной памяти и памяти сборщика мусора на исполнителях. Без хорошего понимания объема данных и доступных вычислительных ресурсов этот подход будет невозможно применить.

Основная особенность этого подхода заключается в том, что каждый исполнитель сможет самостоятельно загружать данные в ОЗУ дата-ноды, на которой он находится, с любой другой машины в кластере, без участия машины-драйвера, что является своего рода «узким местом». С другой стороны, будет загружен блок данных, который действительно нужен в данный момент конкретному исполнителю для тех партиций, на которых он выполняет вычисления, что позволяет своевременно загружать в память часть набора данных, а не целиком весь набор. В некотором смысле это имитация поведения «объединение слиянием», но в нашем случае нет необходимости сортировать строки, что также дает серьезный выигрыш в скорости выполнения.

Алгоритм предлагаемого метода и результаты экспериментов

Алгоритмически предложенный метод реализуется в два этапа. На первом этапе (рис. 7) реорганизуется хранение первого набора данных (RDD1) параллельным чтением строк из одних файлов и записью в другие. Синхронизация доступа к файлам обеспечивается файловой системой HDFS. Оригинальные данные не изменяются, только сортируются. Попадание определенной записи в файл выбирается на основе критерия объединения. Данная операция выполняется без промежуточного этапа перемещения данных между стадиями map и reduce.

На втором этапе (рис. 8) выполняется репартиционирование на логическом уровне также



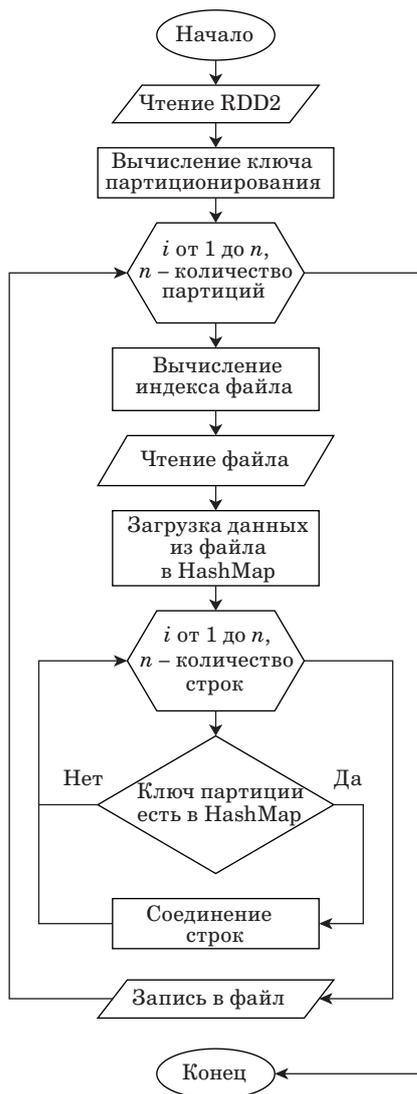
■ **Рис. 7.** Алгоритм первого этапа
 ■ **Fig. 7.** Algorithm of the first stage

на основе заданного критерия объединения без физического перемещения данных по кластеру. Далее для каждой партиции данных параллельно выполняются чтение и загрузка в память файлов первого набора RDD1, где уже локально производится операция сравнения ключей строк, последующее соединение и запись результата.

Разработка во фреймворке Spark доступна на языках Python и R, имеющих большую популярность среди аналитиков. На языках Java и Scala написано ядро фреймворка Spark, поэтому они являются «родными» для него. В отличие от других, язык программирования Scala для Spark предоставляет доступ к новейшим и наиболее важным функциям. Поэтому для реализации предложенного алгоритма выбран язык Scala.

Экспериментальные исследования разработанного метода и его алгоритмической реализации проводились с целью определить время решения задачи объединения неструктурированных текстовых данных объемом от 100 ГБ до 7 ТБ.

В эксперименте были задействованы текстовые данные, сжатые с помощью алгоритма GZIP, 128 программ-исполнителей с шестью ядрами и 16 ГБ памяти и программа драйвера с 8 ГБ памяти. Исходные данные делились равномерно по ключу партиционирования на 3500 партиций. Время объединения данных измерялось стандартной программой spark.time(). Каждое измерение проводилось пять раз, а время выполнения усреднялось.



■ **Рис. 8.** Алгоритм второго этапа
 ■ **Fig. 8.** Algorithm of the second stage

Результаты исследования предложенного в работе метода объединения в сравнении со стандартным Sort-Merge объединением, используемым в Spark SQL [20], представлены на рис. 9.

По результатам проведенных исследований можно сделать вывод, что предложенный метод и его алгоритмическая реализация позволяют решать задачу объединения неструктурированных текстовых данных объемом 2 ТБ на 37 % быстрее, а для данных объемом 7 ТБ примерно на 47 %.

Заключение

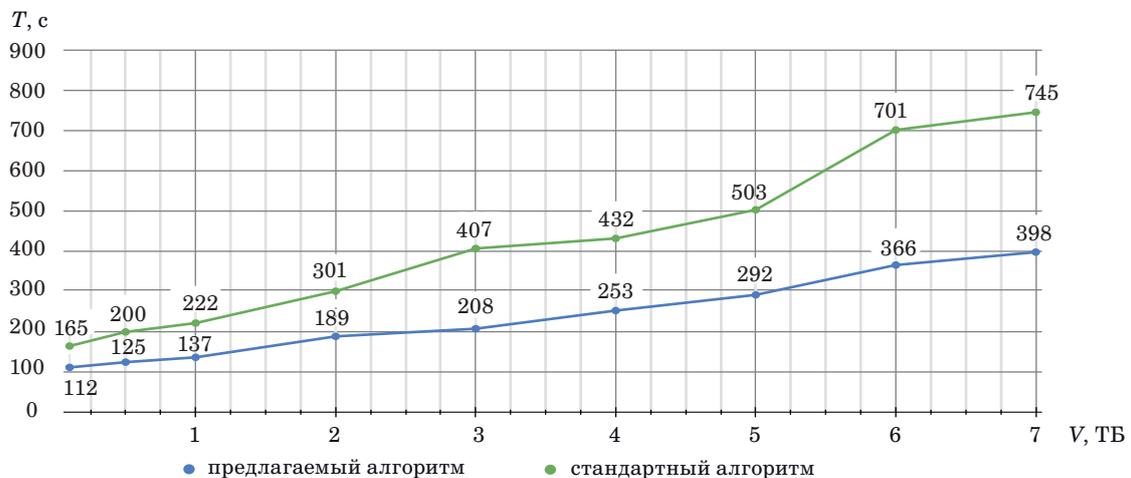
Разработка быстрых и надежных методов обработки больших данных в распределенных системах является актуальной задачей, а имеющиеся методы недостаточно эффективны. Среди них технология Spark является наиболее эффективной с точки зрения скорости вычислений.

Одним из узких мест в кластерных системах являются задачи сортировки и перемещения данных внутри кластера, увеличивающие нагрузку на сеть.

Другой важной задачей при анализе данных является их объединение, методы которого, включая технологию Spark, на сегодняшний момент по критерию времени выполнения недостаточно эффективны.

Предлагаемый метод и его алгоритмическая разработка в сравнении с технологией Spark эффективнее при решении задачи объединения данных на 37–47 %.

Дальнейшее развитие предложенного метода предполагает его модификацию для объединений разных типов данных, построение оптимальных алгоритмов распределения наборов данных по партициям с целью минимизировать их неравномерное разделение.



■ **Рис. 9.** Время выполнения расчета предложенным и стандартным методами
 ■ **Fig. 9.** Time to complete the calculation by the proposed and standard methods

Литература

1. **Yoshimi M., Oge Y., Yoshinaga T.** Pipelined parallel join and its FPGA-based acceleration. *Proc. of the ACM Transactions on Reconfigurable Technology and Systems*, 2007, vol. 10, iss. 4, no. 28, pp 1–28. doi:10.1145/3079759
2. **Zhang X., Chen L., Wang M.** Efficient multi-way theta-join processing using MapReduce. *Proc. of the VLDB Endowment*, 2012, vol. 5, no. 11, pp. 1184–1195. doi:10.14778/2350229.2350238
3. **Saleem M., Potocki A., Soru T., Hartig O., Axel-Cyrille Ngonga Ngomo.** CostFed: Cost-based query optimization for SPARQL endpoint federation. *Proc. of the 14th Intern. Conf. on Semantic Systems*, 2018, Austria, pp. 163–174. doi:10.1016/j.procs.2018.09.016
4. **Sudhaka M., Satheesh N., Balu S., Reddy A. M., Murugan G.** Optimizing joins in a Map-Reduce for data storage and retrieval performance analysis of query processing in HDFS for big data. *Intern. Journal of Advanced Trends in Computer Science and Engineering*, 2019, pp. 2062–2087. doi:10.30534/ijatcse/2019/33852019
5. **Elkawkagy M., Elbeh H.** High performance hadoop distributed file system. *Intern. Journal of Networked and Distributed Computing*, 2020, vol. 8, iss. 3, pp. 119–123. doi:10.2991/ijndc.k.200515.007
6. **Maitrey S., Jha C. K.** MapReduce: Simplified data analysis of big data. *Procedia Computer Science*, 2015, vol. 57, pp. 563–571. doi:10.1016/j.procs.2015.07.392
7. **Malleswari T. Y. J. N., Vadivu G.** MapReduce: A technical review. *Indian Journal of Science and Technology*, 2016, vol. 9, iss. 1, pp. 1–6. doi:10.17485/ijst/2016/v9i1/78964
8. **Shim K.** MapReduce algorithms for big data analysis. *Proc. of the VLDB Endowment*, 2012, vol. 5, no. 12, pp. 2016–2017. doi:10.14778/2367502.2367563
9. **Al-Badarnah A. F., Rababa S. A.** An analysis of two-way equi-join algorithms under MapReduce. *Journal of King Saud University “Computer and Information Sciences”*, 2022, vol. 34, iss. 4, pp. 1074–1085. doi:10.3233/JIFS-201220
10. **Gavagsaz E., Rezaee A., Javadi H. H. S.** Load balancing in join algorithms for skewed data in Map-Reduce systems. *The Journal of Supercomputing*, 2019, vol. 75, pp. 228–254. doi:10.1145/3335484.3335495
11. **Kwon Y. C., Balazinska M., Howe B., Rolia J.** A study of skew in mapreduce applications. *Proc. of the 5th Open Cirrus Summit*, 2011, Moscow.
12. **Gousios G.** Big data software analytics with Apache Spark. *Proc. of the 40th Intern. Conf. on Software Engineering*, 2018, pp. 542–543. doi:10.1145/3183440.3183458
13. **Karau H., Warren R.** *High Performance Spark: Best Practices for Scaling and Optimizing Apache Spark*. O’Reilly, 2017. 342 p.
14. **Bruno N., Kwon Y., Wu M.** Advanced join strategies for large-scale distributed computation. *Proc. of the VLDB Endowment*, 2014, vol. 7, no. 13, pp. 1484–1495. doi:10.14778/2733004.2733020
15. **Saitoh M., Elsayed E. A., Van Chu T., Mashimo S., Kise K.** A high-performance and cost-effective hardware merge sorter without feedback datapath. *IEEE 26th Annual Intern. Symp. on Field-Programmable Custom Computing Machines (FCCM)*, 2018, pp. 197–204. doi:10.1587/transfun.2021EAL2103
16. **Papaphilippou P., Brooks C., Luk W.** FLiMS: Fast lightweight merge sorter. *Proc. of the 29th Intern. Conf. on Field-Programmable Technology (FPT)*, 2018, pp. 78–85. doi:10.48550/arXiv.2112.05607
17. **Papaphilippou P., Pirk H., Luk W.** Accelerating the merge phase of sort-merge join. *Proc. of the 29th Intern. Conf. on Field-Programmable Logic and Applications (FPL)*, 2019, pp. 100–105. doi:10.1109/FPL.2019.00025
18. **Bandle M., Giceva J., Neumann T.** To partition, or not to partition, that is the join question in a real system. *Proc. of the 2021 Intern. Conf. on Management of Data*, 2021, pp. 168–180. doi:10.1145/3448016.3452831
19. **Memarzia P., Ray S., Bhavsar V.** On improving data skew resilience in main-memory hash joins. *Proc. of the 22nd Intern. Database Engineering & Applications Symp.*, 2018, pp. 226–235. doi:10.1145/3216122.3216156
20. **Armbrust M., Xin R. S., Lian C., Huai Y., Liu D., Bradley J. K., Meng X., Kaftan T., Franklin M. J., Ghodsi A., Zaharia M.** Spark SQL: Relational data processing in spark. *Proc. of the Intern. Conf. on Management of Data (SIGMOD’15)*, 2015, Australia, pp. 1383–1394. doi:10.1145/2723372.2742797

UDC 004.657

doi:10.31799/1684-8853-2022-5-2-11

EDN: AAGGGR

A method to accelerate the joining of distributed datasets by a given criterionY. S. Tyryshkina^a, Post-Graduate Student, orcid.org/0000-0003-4814-4874S. R. Tumkovskiy^a, Dr. Sc., Tech., Associate Professor, orcid.org/0000-0002-1647-2171, STumkovskiy@hse.ru^aNational Research University «Higher School of Economics», 20, Myasnitskaya St., 101000, Moscow, Russian Federation

Introduction: Rapidly growing volumes of information pose new challenges to modern data analysis technologies. One of the most common related operations in analytics is the joins of datasets. A join is a highly resource-intensive operation which is difficult to scale and improve in terms of resource efficiency in distributed databases or systems based on the MapReduce paradigm. **Purpose:** To develop a method to accelerate the integration of datasets in distributed systems. **Results:** The architecture of Apache Spark, features of distributed computing based on MapReduce are examined, typical methods of joining datasets are analyzed, and the main ways to speed up the operation of joining data are considered. We propose a method to speed up a special case of joins implemented in Apache Spark with the use of partitioning and partial transfer of datasets to cluster computing nodes. The method is realized to simultaneously harness the benefits of merge and broadcast joins. The presented experimental data demonstrate that the larger the volume of input data the more efficiently the method speeds up the operation of joining data in comparison with standard methods. Thus, for 2 TB compressed data, acceleration up to ~37% was obtained in comparison with standard Spark SQL.

Keywords – distributed computing, MapReduce, Apache Spark, joining data.

For citation: Tyryshkina Y. S., Tumkovskiy S. R. A method to accelerate the joining of distributed datasets by a given criterion. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 5, pp. 2–11 (In Russian). doi:10.31799/1684-8853-2022-5-2-11, EDN: AAGGGR

References

1. Yoshimi M., Oge Y., Yoshinaga T. Pipelined parallel join and its FPGA-based acceleration. *Proc. of the ACM Transactions on Reconfigurable Technology and Systems*, 2007, vol. 10, iss. 4, no. 28, pp 1–28. doi:10.1145/3079759
2. Zhang X., Chen L., Wang M. Efficient multi-way theta-join processing using MapReduce. *Proc. of the VLDB Endowment*, 2012, vol. 5, no. 11, pp. 1184–1195. doi:10.14778/2350229.2350238
3. Saleem M., Potocki A., Soru T., Hartig O., Axel-Cyrille Ngonga Ngomo. CostFed: Cost-based query optimization for SPARQL endpoint federation. *Proc. of the 14th Intern. Conf. on Semantic Systems*, 2018, Austria, pp. 163–174. doi:10.1016/j.procs.2018.09.016
4. Sudhaka M., Satheesh N., Balu S., Reddy A. M., Murugan G. Optimizing joins in a Map-Reduce for data storage and retrieval performance analysis of query processing in HDFS for big data. *Intern. Journal of Advanced Trends in Computer Science and Engineering*, 2019, pp. 2062–2087. doi:10.30534/ijatcse/2019/33852019
5. Elkawkagy M., Elbeh H. High performance hadoop distributed file system. *Intern. Journal of Networked and Distributed Computing*, 2020, vol. 8, iss. 3, pp. 119–123. doi:10.2991/ijndc.k.200515.007
6. Maitrey S., Jha C. K. MapReduce: Simplified data analysis of big data. *Procedia Computer Science*, 2015, vol. 57, pp. 563–571. doi:10.1016/j.procs.2015.07.392
7. Malleswari T. Y. J. N., Vadivu G. MapReduce: A technical review. *Indian Journal of Science and Technology*, 2016, vol. 9, iss. 1, pp. 1–6. doi:10.17485/ijst/2016/v9i1/78964
8. Shim K. MapReduce algorithms for big data analysis. *Proc. of the VLDB Endowment*, 2012, vol. 5, no. 12, pp. 2016–2017. doi:10.14778/2367502.2367563
9. Al-Badarneh A. F., Rababa S. A. An analysis of two-way equi-join algorithms under MapReduce. *Journal of King Saud University “Computer and Information Sciences”*, 2022, vol. 34, iss. 4, pp. 1074–1085. doi:10.3233/JIFS-201220
10. Gavagsaz E., Rezaee A., Javadi H. H. S. Load balancing in join algorithms for skewed data in MapReduce systems. *The Journal of Supercomputing*, 2019, vol. 75, pp. 228–254. doi:10.1145/3335484.3335495
11. Kwon Y. C., Balazinska M., Howe B., Rolia J. A study of skew in mapreduce applications. *Proc. of the 5th Open Cirrus Summit*, 2011, Moscow.
12. Gousios G. Big data software analytics with Apache Spark. *Proc. of the 40th Intern. Conf. on Software Engineering*, 2018, pp. 542–543. doi:10.1145/3183440.3183458
13. Karau H., Warren R. *High Performance Spark: Best Practices for Scaling and Optimizing Apache Spark*. O’Reilly, 2017. 342 p.
14. Bruno N., Kwon Y., Wu M. Advanced join strategies for large-scale distributed computation. *Proc. of the VLDB Endowment*, 2014, vol. 7, no. 13, pp. 1484–1495. doi:10.14778/2733004.2733020
15. Saitoh M., Elsayed E. A., Van Chu T., Mashimo S., Kise K. A high-performance and cost-effective hardware merge sorter without feedback datapath. *IEEE 26th Annual Intern. Symp. on Field-Programmable Custom Computing Machines (FCCM)*, 2018, pp. 197–204. doi:10.1587/transfun.2021EAL2103
16. Papaphilippou P., Brooks C., Luk W. FLiMS: Fast lightweight merge sorter. *Proc. of the 29th Intern. Conf. on Field-Programmable Technology (FPT)*, 2018, pp. 78–85. doi:10.48550/arXiv.2112.05607
17. Papaphilippou P., Pirk H., Luk W. Accelerating the merge phase of sort-merge join. *Proc. of the 29th Intern. Conf. on Field-Programmable Logic and Applications (FPL)*, 2019, pp. 100–105. doi:10.1109/FPL.2019.00025
18. Bandle M., Giceva J., Neumann T. To partition, or not to partition, that is the join question in a real system. *Proc. of the 2021 Intern. Conf. on Management of Data*, 2021, pp. 168–180. doi:10.1145/3448016.3452831
19. Memarzia P., Ray S., Bhavsar V. On improving data skew resilience in main-memory hash joins. *Proc. of the 22nd Intern. Database Engineering & Applications Symp.*, 2018, pp. 226–235. doi:10.1145/3216122.3216156
20. Armbrust M., Xin R. S., Lian C., Huai Y., Liu D., Bradley J. K., Meng X., Kaftan T., Franklin M. J., Ghodsi A., Zaharia M. Spark SQL: Relational data processing in spark. *Proc. of the Intern. Conf. on Management of Data (SIGMOD’15)*, 2015, Australia, pp. 1383–1394. doi:10.1145/2723372.2742797



Аналитический обзор методов автоматического распознавания вовлеченности пользователя в виртуальную коммуникацию

А. А. Двойникова^а, младший научный сотрудник, orcid.org/0000-0001-8047-6639

И. А. Кагиров^а, научный сотрудник, orcid.org/0000-0003-1196-1117

А. А. Карпов^а, доктор техн. наук, профессор, orcid.org/0000-0003-3424-652X, karpov@iias.spb.su

^аСанкт-Петербургский Федеральный исследовательский центр РАН, 14-я линия В. О., 39, Санкт-Петербург, 191718, РФ

Введение: решение автоматическими средствами задачи распознавания и оценивания степени вовлеченности пользователя в процесс человеко-машинного взаимодействия или телекоммуникации является актуальным в области компьютерного распознавания состояний человека. Это необходимо для проектирования приложений дистанционного обучения, бизнеса и развлечений. **Цель:** провести сравнительный анализ существующего информационного обеспечения и методов в области автоматического распознавания и оценивания вовлеченности пользователя в процесс человеко-машинного взаимодействия или виртуальной коммуникации и обосновать методологию построения корпуса данных, основанного на идее многомодальной коммуникации. **Результаты:** выполненный анализ работ показал, что в большей части известных корпусов (*data sets*) отсутствуют данные по онлайн-коммуникации в естественных условиях и не всегда учитываются разные модальности взаимодействия в системе «человек – машина – человек». Для многоуровневой классификации, определяющей степень вовлеченности, важны, помимо видеомодальности, также акустическая и текстовая информация. Перспективным оказывается учет «языка тела» пользователя: мимика, движения тела, рук и головы. Кроме того, для правильной оценки вовлеченности в корпусе данных должна содержаться метаинформация по психоэмоциональным состояниям коммуникантов. Наиболее перспективным подходом для автоматического оценивания степени вовлеченности следует признать нейросетевой. **Практическая значимость:** на основании полученных аналитических выводов планируется разработка оригинальной программной системы для автоматического распознавания вовлеченности и формирование набора данных для обучения ее вероятностных моделей. Также представлен обзор позволяет сформулировать основные требования к подобным системам и является вкладом в решение задачи автоматического распознавания психоэмоциональных состояний человека. **Обсуждение:** анализ работ позволяет сделать вывод, что понятие вовлеченности в контексте распознавания эмоций отличается от распространенного в психологии. Вовлеченность пользователя (коммуниканта) в информационно-коммуникационной сфере – проявление различной степени эмоциональной, когнитивной и поведенческой составляющих психической активности человека в процессе взаимодействия с собеседником или компьютерной системой, имеющее динамический характер.

Ключевые слова – вовлеченность пользователя, информационное обеспечение, автоматические системы распознавания вовлеченности, многомодальность, искусственные нейронные сети.

Для цитирования: Двойникова А. А., Кагиров И. А., Карпов А. А. Аналитический обзор методов автоматического распознавания вовлеченности пользователя в виртуальную коммуникацию. *Информационно-управляющие системы*, 2022, № 5, с. 12–22. doi:10.31799/1684-8853-2022-5-12-22, EDN: CXBRCS

For citation: Dvoynikova A. A., Kagirow I. A., Karpov A. A. Analytical review of methods for automatic detection of user engagement in virtual communication. *Informatsionno-upravliayushchie sistemy* [Information and Control Systems], 2022, no. 5, pp. 12–22 (In Russian). doi:10.31799/1684-8853-2022-5-12-22, EDN: CXBRCS

Введение

В современном мире средства онлайн-коммуникации уже давно и прочно являются одним из основных способов межличностного взаимодействия. В период пандемии COVID (с 2020 г.) актуальность дистанционных технологий только возросла. Не в последнюю очередь это касается учебных процессов, организации рабочих совещаний и телеконференций [1]. Другой существенной составляющей онлайн-коммуникации является повседневное общение и развлечение (прежде всего, массовые многопользовательские игры, зачастую требующие от участников обсуждения командных действий).

Перенос многих форм социального взаимодействия в дистанционный формат порождает ряд новых проблем, одной из которых является низкая вовлеченность (Engagement) одного или нескольких участников в коммуникативный процесс (в широком смысле этого слова). Отсутствие живого контакта между участниками онлайн-мероприятий (эффект присутствия), во-первых, может способствовать рассеянности внимания собеседников и, во-вторых, лишает коммуникантов возможности прямо оценивать уровень заинтересованности собеседника.

Эта проблема актуальна, прежде всего, для онлайн-обучения. Если вовлеченность учащихся

в учебный процесс и раньше оказывалась в центре внимания научного сообщества, то сегодня разработку методов и средств объективного измерения степени вовлеченности в образовательный процесс можно признать необходимостью [2, 3].

Вовлеченности как синониму осознанной активной внимательности посвящен целый ряд психологических исследований [4, 5], результаты которых находят практическое применение во многих областях человеческой деятельности. Например, людям, работающим в сфере управления, иногда критически важно удерживать свое внимание на рабочих процессах (мониторинг, визуальный контроль и принятие решений на основе видеоданных). В чрезвычайных ситуациях оператору необходимо быстро принять правильное решение, а этого можно достичь только при вовлеченности (в другой терминологии — активном включении) оператора в рабочий процесс. Поэтому психологи совместно с техническими специалистами разрабатывают различные тренажеры для тренировки внимания [4, 5]. Системы управления и предоставления информации должны не просто передавать визуальные данные, но и оценивать, в том числе, эмоциональное состояние людей, облегчая процесс принятия решений [6].

Главной целью настоящей статьи является сравнительный аналитический обзор основных методов распознавания вовлеченности участников виртуальной коммуникации, в последнее время появившихся в контексте исследований по машинному обучению. Под виртуальной коммуникацией здесь подразумевается как общение между собеседниками посредством телеконференций, так и взаимодействие пользователя с системой (просмотр видео, онлайн-игры и пр.). Другой важной задачей оказывается выполнение аналитического обзора существующего информационного обеспечения, созданного для распознавания вовлеченности участников.

Термин «вовлеченность» был введен в научный оборот в работе [7], и, начиная с конца XX в., он активно используется в научном сообществе, в первую очередь среди специалистов по психологии, педагогике и социологии. Тем не менее во всех этих дисциплинах вовлеченность трактуется по-разному, и на сегодня не существует консенсуса относительно его научного содержания [8]. Можно констатировать, что понятие вовлеченности прочно вошло в систему научного знания, однако эксплицитное раскрытие этого термина является предметом научных дискуссий [9]. Вовлеченность является сложным психологическим и социальным феноменом; различные исследователи понимают этот термин с разных позиций [10]. Необходимо подчеркнуть при этом, что разнообразие мнений о вовлеченности на-

прямую проистекает из сложной природы этого феномена.

В работе [11] выделяется три основных компонента социальной установки человека: когнитивный (познавательный) [12], аффективный (эмоциональный) [13, 14] и поведенческий [15, 16]. В зависимости от того, в каком компоненте применяется термин «вовлеченность», его определение может несколько отличаться [17].

Также существуют работы, в которых проводятся исследования корреляции всех трех аспектов вовлеченности: познавательного, эмоционального и поведенческого. Так, в статье [18] указывается, что наиболее глубокая погруженность в работу отмечается возросшей когнитивной составляющей, что подразумевает у индивида процесс осмысления задачи. Исследование [19] демонстрирует взаимосвязь между когнитивной вовлеченностью и обучаемостью студентов. В работе [13] экспериментальным путем доказывается связь между эмоциональной и поведенческой вовлеченностью.

В контексте автоматического распознавания вовлеченности следует учитывать тот факт, что автоматическими методами можно анализировать только внешнее проявление, а не истинную вовлеченность человека. Вовлеченность в области автоматического анализа — это проявление различной степени эмоциональной, когнитивной и поведенческой составляющих человека в процессе взаимодействия с собеседником или компьютерной системой, имеющее динамический характер. По различным уровням интенсивности ряд авторов выделяет вовлеченность бинарную (вовлечен/не вовлечен) [20, 21], трехуровневую (не вовлечен, вовлечен формально, сильно вовлечен) [22], четырехуровневую (очень низкую, низкую, высокую и очень высокую) [23–25] и даже десятиуровневую [26]. Помимо категориального разделения степени вовлеченности, также активно применяется и непрерывная оценка вовлеченности [27, 28].

Обзор информационного обеспечения для распознавания вовлеченности

Для того чтобы разработать систему автоматического распознавания вовлеченности, необходимо построить модель с учетом реальных данных. Все существующие корпуса, содержащие разметку данных по вовлеченности, можно разделить на две группы по принципу взаимодействия участников. К первой группе относятся корпуса, при записи которых участники коммуницировали друг с другом, такие как NoXi [27], MEDICA [28], MHHRI [29], RECOLA [20], Emotion Miner Data Corpus (EMDC) и Sümer [2]. Другая группа пред-

ставляет собой корпуса, содержащие сценарии, в которых участники взаимодействуют с компьютерной системой, например просматривают обучающие видео или играют в онлайн-игры. К таким корпусам относятся Engagement Recognition dataset [21], EngageWild [23], DAiSEE [24], корпуса авторов Kamath [22], Whitehill [25], Psaltis [29]. В табл. 1 содержится сравнительный обзор корпусов, применяемых для анализа вовлеченности.

Как видно из таблицы, существует несколько представительных корпусов для автоматического распознавания вовлеченности. Стоит отметить, что большинство рассмотренных корпусов включают в себя только видеоданные. Это обусловлено тем, что сценарии для записи подразумевают только просмотр участниками видеосюжетов или участие в онлайн-играх. В таких корпусах представлена вовлеченность участников во взаимодействие с компьютерной системой. В многомодальных корпусах содержатся данные по вовлеченности участников в процесс коммуникации друг с другом. Проявление вовлеченности в диалоге с собеседником зависит не только от собственного интереса к теме разговора, но и от черт личности собеседника, его коммуникативных на-

выков, а также от проявления эмоционального состояния обоих коммуникантов.

Наибольшая вариативность информантов среди одномодальных корпусов представлена в корпусах Sümer [2] (128 чел.) и DAiSEE [24] (112 чел.). При этом корпус EMDC имеет наибольший объем видеоданных (140 ч) по сравнению со всеми рассмотренными корпусами. Корпус NoXi [27] содержит в себе данные большого количества участников (87 чел.), которые общались между собой на различные темы на семи языках. Исходя из этого можно сказать, что NoXi является наиболее репрезентативным корпусом для обучения многомодальных систем автоматического распознавания вовлеченности.

Еще одним немаловажным фактором при анализе данных для распознавания вовлеченности является учет эффекта Хоторна (эффект наблюдателя) [30] при записи участников эксперимента. Эффект Хоторна представляет собой реакцию – изменение поведения человека в ответ на осознание того, что за ним наблюдают со стороны. Во всех рассмотренных корпусах участники экспериментов знали о предстоящей записи, поэтому могли вести себя не вполне естественно.

■ **Таблица 1.** Сравнительная характеристика корпусов для автоматического анализа вовлеченности участников
 ■ **Table 1.** Comparative characteristics of corpora collected for automatic analysis of communicant engagement

Корпус	Модальность				Объем, ч	Разметка данных
	В	А	Т	Ф		
NoXi (https://nox.aria-agent.eu/) [27]	+	+	+	-	25,3	Вовлеченность (непрерывная), эмоции, жесты, мимика
MEDICA [28]	+	+	+	-	1,1	Вовлеченность (непрерывная), внимательность, стресс, эмоции, нерешительность
MNHRI [26]	+	+	-	+	4,2	Вовлеченность (10 уровней), характеристики личности
RECOLA [20]	+	+	-	+	3,8	Вовлеченность (бинарная), согласие, доминирование, производительность, взаимопонимание
Emotion Miner Data Corpus (EMDC) (https://www.prweb.com/releases/2018/03/prweb15339730.htm)	+	+	-	-	140	Разговор (вовлеченность), эмоции, психические состояния и поведение, личность и ситуация
Корпус Sümer [2]	+	-	-	-	47	Вовлеченность (6 уровней)
Корпус Whitehill [25]	+	-	-	-	22,6	Вовлеченность (4 уровня)
DAiSEE (https://www.iith.ac.in/~daisee-dataset/) [24]	+	-	-	-	25,2	Вовлеченность (4 уровня), скука, замешательство, разочарование
Engagement Recognition dataset [21]	+	-	-	-	20	Вовлеченность (3 уровня), эмоции
Корпус Psaltis [29]	+	-	-	-	12	Вовлеченность (бинарная)
EngageWild (https://www.sites.google.com/view/emotiw2020) [23]	+	-	-	-	8,5	Вовлеченность (4 уровня)
Корпус Kamath (https://github.com/edrishti/wacv2016) [22]	+	-	-	-	4408 кадров	Вовлеченность (3 уровня)

Примечание: В – видео; А – аудио; Т – текст; Ф – физиологические сигналы.

В корпусах DAiSEE и Kamath [22] использовалась веб-камера, которая позволяла приблизить условия эксперимента к естественным, потому что веб-камера незаметна и является довольно привычным предметом современной обстановки.

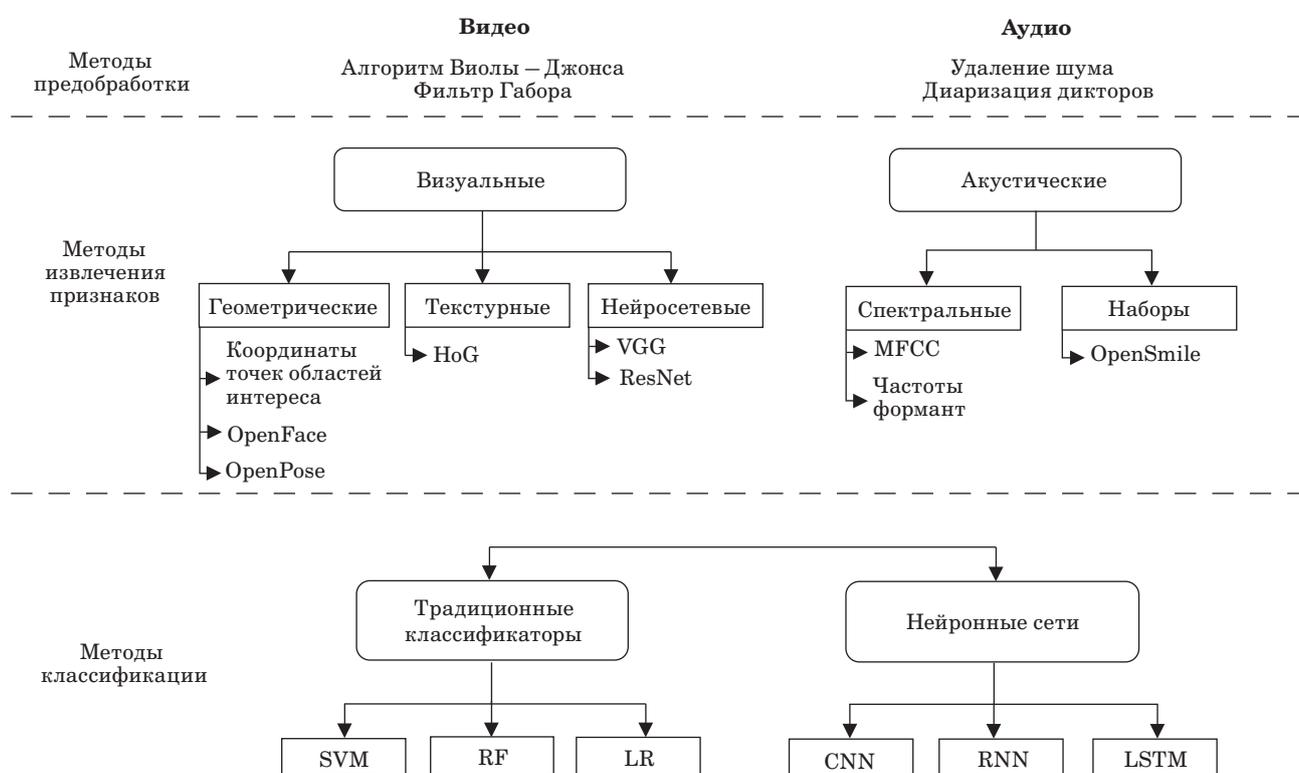
Обзор методов для распознавания вовлеченности

Большинство автоматических систем распознавания вовлеченности участников виртуальной коммуникации или взаимодействия с компьютерной системой основываются на анализе репрезентативных данных. Наличие текстовой информации в высказывании диктора может символизировать об однозначной вовлеченности диктора. Визуальные и акустические характеристики человека могут передавать как вовлеченность человека в виртуальную коммуникацию, так и ее отсутствие. Поэтому наиболее репрезентативными модальностями являются видео- и аудиоданные. При анализе видеоданных необходимо в первую очередь детектировать область лица, а также жесты рук, головы, позы тела и пр. Затем проводится извлечение визуальных и акустических признаков, которые в последующем подаются на вход классификатора, и на выходе

даются вероятностные предсказания степени вовлеченности участников. Алгоритм распознавания вовлеченности с применяемыми методами предобработки данных, извлечения признаков и классификации представлен на рисунке.

Методы предобработки. Для выделения репрезентативной информации из видеоданных необходимо детектировать графические области интереса, которые представляют собой границы областей лица, тела, рук и пр. Алгоритм Виолы – Джонса наиболее часто используется для детектирования области лица в задаче распознавания вовлеченности [22]. К изображениям детектированных областей интереса также может применяться нормализация, например с помощью фильтра Габора, который преобразует изображения в оттенки серого [25]. Данные методы позволяют выделить релевантную информацию из изображений для анализа мимики лица и жестов.

Методы извлечения признаков. Все визуальные признаки, используемые для анализа вовлеченности, можно разделить на три группы: геометрические, текстурные и нейросетевые. Геометрические признаки используются для определения области положения губ [31], зрачков [31], позы тела [32], глаз [23, 33], головы [23], [32, 33]. Благодаря этому можно определить, откры-



- Систематизация методов распознавания вовлеченности пользователей в виртуальную коммуникацию
- Systematization of methods for recognition of user engagement into virtual communication

ты или закрыты глаза у собеседника, смотрит ли диктор в монитор или отвлекается, по положению губ можно сказать о том, произносит ли речь диктор, или анализировать эмоции (например, улыбка показывает радость). Также по жестам тела можно определить степень вовлеченности диктора, например, если рука подпирает голову, то это говорит о низкой заинтересованности. OpenFace и OpenPose – программные инструментари, позволяющие находить координаты лица и частей тела соответственно. Помимо этого, OpenFace также выделяет единицы действия лица, например, закрыты глаза или открыты. OpenFace и OpenPose применялся в работах [2, 34–38] для распознавания вовлеченности собеседников.

Наиболее часто используемым методом извлечения текстурных признаков для распознавания вовлеченности [21, 22] является гистограмма ориентированных градиентов (Histogram of Oriented Gradients, HoG). HoG позволяет выделять контурную информацию объектов, удаляя при этом фоновый шум. Некоторые исследователи используют предобученные нейронные сети, такие как VGG [32, 38], ResNet (Residual Neural Network) [31], для извлечения визуальных признаков. Однако нейросетевые признаки не несут в себе физической информации об объектах, но зачастую позволяют достигать высокого результата классификации уровня вовлеченности.

Для распознавания вовлеченности в основном используют спектральные акустические признаки, такие как мел-частотные кепстральные коэффициенты (Mel-Frequency Cepstral Coefficients, MFCC), частоты формант и пр. [26, 28]. Данные характеристики отражают физическую составляющую речевого сигнала и зависят от особенностей работы артикуляционных органов каждого человека. Также одним из используемых наборов признаков для распознавания вовлеченности [31] является OpenSmile, который включает в себя 65 низкоуровневых дескрипторов.

Методы классификации. Существует множество методов классификации вовлеченности, их можно разделить на две группы: традиционные детерминированные методы классификации и искусственные нейронные сети (см. рисунок).

Эффективным традиционным методом распознавания вовлеченности является метод опорных векторов (Support Vector Machine, SVM) [22, 25, 26, 29]. Авторы работы [26] применяли SVM для бинарной классификации вовлеченности на основе анализа различных типов данных: видео, аудио и физиологических сигналов. Многие исследователи использовали SVM для классификации только видеоданных по кадрам [22, 25,

29]. В работе [22] автоматическая система распознавала три степени вовлеченности, используя SVM с обучением на нескольких ядрах. Другой метод традиционных классификаторов, который используется для распознавания вовлеченности, – случайный лес (Random Forest, RF) [2].

При наличии большого объема данных наиболее эффективными методами машинной классификации являются нейронные сети. Для анализа видеоданных можно выделить две основные группы нейронных сетей: прямого распространения, например сверточные (Convolutional Neural Network, CNN), и с обратными связями, например рекуррентные (Recurrent Neural Network, RNN). В работе [21] для бинарной классификации вовлеченности авторы используют архитектуру нейросети VGG-B, предварительно обученную на эмоциональном корпусе данных. Авторы статьи [35] для распознавания вовлеченности предлагают подход интегрального распознавания (End-to-End), основанный на использовании архитектур нейронных сетей ResNet и Temporal Convolutional Network (TCN).

Рекуррентные нейронные сети также являются эффективными в задачах распознавания вовлеченности участников виртуальной коммуникации. В работе [24] использовалась архитектура нейросети InceptionNet V3 [39] для кадрового анализа видеоданных и нейронная сеть Long-Term Recurrent Convolutional Network (LRCN) для классификации всего видео целиком. Авторы статьи [36] применяют нейронную сеть с управляемыми рекуррентными нейронами (Gated Recurrent Unit, GRU). Также для анализа вовлеченности используются нейронные сети с долгой краткосрочной памятью (Long Short-Term Memory, LSTM) [23, 34]. В статье [37] используются несколько LSTM-сетей: для анализа диктора, собеседника и обоих коммуникантов.

Для повышения уровня точности распознавания вовлеченности применяется объединение модальностей. В работе [32] анализируются видео- и аудиоданные с помощью байесовской сети (Bayesian Network, BN), а также RNN, а в работе [31] – с помощью логистической регрессии (Logistic Regression, LR).

Анализ исследований по распознаванию вовлеченности

В настоящее время существует ряд коммерческих систем для анализа вовлеченности персонала компаний в их трудовую деятельность. Наиболее известными являются системы TalentTech (<https://talenttech.ru/engagement/>), Happy Job (<https://happy-job.ru/>), Gallup ([16](https://www.gallup.com/workplace/229424/employee-</p>
</div>
<div data-bbox=)

engagement.aspx). Однако данные системы анализируют вовлеченность сотрудников компании при помощи методологии, основанной на опросниках [40–42]. Коммерческих систем, основанных на анализе аудиовизуальной информации проявления вовлеченности, на данный момент не существует. Именно поэтому далее в настоящей статье рассматриваются некоммерческие автоматические системы, работающие с аудио и визуальными данными.

Характеристики систем для автоматического распознавания вовлеченности пользователя приведены в табл. 2. Для оценки эффективности применяются коэффициент корреляции Пирсона (Pearson's Correlation Coefficient, PCC), среднеквадратическая ошибка (Mean Squared Error, MSE), а также точность (Precision) и F-мера (F-score).

Как показывает обзор существующих работ, в настоящее время чаще всего система проводит анализ видеоданных, а не аудио и текста. Это обусловлено тем, что видеоданные являются более репрезентативными для анализа вовлеченности и в ситуациях, когда взаимодействие человека происходит не с собеседником, а, например, с информационной системой. В таких случаях заинтересованность проявляется только в мимике и жестах. Проявление речевой активности, как правило, говорит о том, что человек вовлечен, поэтому при бинарной классификации вовлеченности нет смысла анализировать акустические и смысловые признаки. Однако при многоуровневой классификации, показывающей интенсивность вовлеченности, анализ аудио и текстовой модальностей может помочь повысить точность распознавания.

Исследований, направленных на анализ физиологических сигналов человека для решения данной задачи, довольно мало. Это связано с тем, что сбор таких данных является достаточно трудоемким процессом, к тому же носимые участниками эксперимента датчики могут мешать естественному поведению человека. Также анализ аудио и текстовой модальностей в области распознавания вовлеченности является недостаточно изученным. Актуальность данной задачи проявляется в ситуациях, когда анализ видеоданных становится невозможным, например, при неработающей камере при виртуальной коммуникации.

Использование нейронных сетей для распознавания вовлеченности показывает более высокую точность по сравнению с применением традиционных методов классификации (см. табл. 2). Одной из причин эффективности нейронных сетей является возможность использовать предварительно обученные модели на других корпусах с большим объемом данных с переносом обучения (Transfer Learning). Некоторые существующие экспериментальные исследования показывают

Использование нейронных сетей для распознавания вовлеченности показывает более высокую точность по сравнению с применением традиционных методов классификации (см. табл. 2). Одной из причин эффективности нейронных сетей является возможность использовать предварительно обученные модели на других корпусах с большим объемом данных с переносом обучения (Transfer Learning). Некоторые существующие экспериментальные исследования показывают

■ **Таблица 2.** Сравнение исследований по автоматическому распознаванию вовлеченности

■ **Table 2.** Comparison of studies on automatic engagement recognition

Метод (работа)	Корпус	Модальность	Методы классификации	Показатель оценки	Значение показателя	
[38]	DAiSEE	B	RNN	Precision	0,39	
[24]			InceptionNet		0,47	
[35]			LRCN		0,58	
[36]			ResNet, TCN		0,64	
[33]	EngageWild		GRU	MSE	0,07	
[34]			TCN		0,08	
[23]			LSTM		PCC	0,06
[37]	Средняя F-score			0,25		
[32]	NoXi		A+B	RNN	PCC	0,60
[32]				BN		0,74
[26]	MHHRI	A	SVM	F-score	0,65	
		B			0,58	
		Φ			0,54	
		A+B			0,56	
		A+Φ			0,59	
		B+Φ			0,52	

эффективность использования предобученных моделей на эмоциональных корпусах для анализа вовлеченности. Это связано со значительной корреляцией между эмоциональным состоянием человека и вовлеченностью как в коммуникациях, так в образовательных и игровых [21, 43–45] взаимодействиях между людьми.

Исходя из экспериментальных исследований [24, 35, 38], можно сделать вывод, что с помощью архитектуры ResNet получается более эффективное распознавание вовлеченности, чем с RNN. На визуальном корпусе DAiSEE авторы работы [35] достигают точности 64 %, используя нейронную сеть ResNet. LSTM [34] позволяет уменьшить MSE относительно GRU [36] и TCN [33] до 0,06 на корпусе EngageWild.

Экспериментальные исследования показывают, что для повышения точности распознавания вовлеченности, помимо анализа мимики, эффективно анализировать поворот головы, жесты и направление взгляда [29, 32]. Также, как правило, объединение нескольких модальностей: видео, аудио, текста, физиологических сигналов – позволяет улучшить точность распознавания вовлеченности. Однако в работе [26] эксперименты показали противоположный результат: одномодальные подходы превзошли многомодальные. Авторы предполагают, что это связано со способом объединения модальностей на уровне принятия решения. Вероятно, объединение на уровне признаков позволит повысить точность одномодального распознавания.

Заключение

Как показывает обзор существующих корпусов, для анализа вовлеченности практически все данные собираются в лабораторных (контролируемых) условиях. В том случае, когда авторы корпусов рассматривают сценарии коммуникации между людьми, общение происходит только между двумя участниками. Ни один из описанных выше корпусов не включает в себя коммуникацию между людьми в группе. К тому же на сегодня все больше коммуникативных актов происходит в онлайн-режиме. Также стоит отметить, что большинство многомодальных корпусов англоязычные, меньшая часть – на французском и немецком языках.

Аналитический обзор существующих методов анализа вовлеченности позволяет сформулировать следующие основные требования к разрабатываемой нами программной системе автоматического распознавания вовлеченности:

1. Многомодальный анализ данных вербальных и невербальных сигналов проявлений вовлеченности коммуникантов.

2. Распознавание вовлеченности с высокой точностью (не менее 70 %).

3. Учет эмоционального состояния коммуникантов.

4. Поддержка офлайн- и онлайн-режимов.

5. Возможность интеграции в существующие системы телеконференций.

6. Использование нейросетевых подходов.

Основываясь на этих выводах, авторы статьи планируют собрать собственный корпус данных, в котором должны быть удовлетворены следующие требования:

1. Многомодальность: корпус должен включать в себя видео, аудио и текстовые данные.

2. Включение сценариев коммуникации группы от двух и более людей.

3. Запись данных в естественных условиях с использованием современных средств телекоммуникации (например, Zoom, Signal, Яндекс, Телемост и т. п.).

4. Вербальная коммуникация дикторов между собой на русском языке.

5. Минимизация эффекта Хоторна при записи данных.

6. Разметка данных по меткам вовлеченности и психоэмоциональных состояний информантов.

Таким образом, мы планируем собрать новый русскоязычный многомодальный корпус, содержащий записи коммуникаций людей в группе. На основе этих данных будет разработана программная система для распознавания вовлеченности с использованием многомодального анализа информации.

Финансовая поддержка

Настоящий обзор выполнен в рамках ведущей научной школы РФ (грант № НШ-17.2022.1.6), а также частично в рамках бюджетной темы (№ FFZF-2022-0005).

Литература

1. Pregowska A., Masztalerz K., Garlińska M., Osial M. A worldwide journey through distance education – from the post office to virtual, augmented and mixed realities, and education during the COVID-19 pandemic. *Education Sciences*, 2021, vol. 11, no. 3, pp. 1–26. doi:10.3390/educsci11030118
2. Sümer Ö., Goldberg P., d’Mello S., Gerjets P., Trautwein U., Kasneci E. Multimodal engagement analysis from facial videos in the classroom. *IEEE Transactions on Affective Computing*, 2021, 16 p. doi:10.1109/TAFFC.2021.3127692

3. **Nkomo L., Daniel B.** Sentiment analysis of student engagement with lecture recording. *TechTrends*, 2021, vol. 65, no. 2, pp. 213–224. doi:10.1007/s11528-020-00563-8
4. **Дозорцев В. М., Назин В. А.** Компьютерные тренажеры как инструмент моделирования операторской деятельности в психологическом эксперименте. *Актуальные проблемы психологии труда, инженерной психологии и эргономики*: тр. Института психологии РАН, 2013, вып. 5, с. 81–103.
5. **Соколов В. Н., Коротеев Г. Л.** Принципы и технологии построения адаптивных обучающих сред. *Актуальные проблемы психологии труда, инженерной психологии и эргономики*: тр. Института психологии РАН, 2013, вып. 5, с. 57–81.
6. **Двойникова А. А., Карпов А. А.** Аналитический обзор подходов к распознаванию тональности русскоязычных текстовых данных. *Информационно-управляющие системы*, 2020, № 4, с. 20–30. doi:10.31799/1684-8853-2020-4-20-30
7. **Kahn W. A.** Psychological conditions of personal engagement and disengagement at work. *The Academy of Management Journal*, 1990, vol. 33, no. 4, pp. 692–724. doi:10.2307/256287
8. **Kelders S. M., van Zyl L. E., Ludden G.** The concept and components of engagement in different domains applied to eHealth: A systematic scoping review. *Frontiers in Psychology*, 2020, vol. 11, Article 926. doi:10.3389/fpsyg.2020.00926
9. **De Vreede T., Andel S., de Vreede G.-J., Spector P. E., Singh V., Padmanabhan B.** What is engagement and how do we measure it? Toward a domain independent definition and scale. *Proc. of the 52nd Hawaii Intern. Conf. on System Sciences (HICSS 2019)*, 2019, pp. 1–10. doi:10.24251/HICSS.2019.092
10. **Calder B. J., Malthouse E. C., Schaedel U.** An experimental study of the relationship between online engagement and advertising effectiveness. *Journal of Interactive Marketing*, 2009, vol. 23, no. 4, pp. 321–331. doi:10.1016/j.intmar.2009.07.002
11. **Smith M.** An approach to the study of the social act. *Psychological Review*, 1942, vol. 49, no. 5, pp. 422–440. doi:10.1037/h0062907
12. **Posner M. I.** Orienting of attention. *Quarterly Journal of Experimental Psychology*, 1980, vol. 32, no. 1, pp. 3–25. doi:10.1080/00335558008248231
13. **Li Y., Lerner R. M.** Interrelations of behavioral, emotional, and cognitive school engagement in high school students. *Journal of Youth and Adolescence*, 2013, vol. 42, no. 1, pp. 20–32. doi:10.1007/s10964-012-9857-5
14. **Truss C., Soane E., Edwards C., Wisdom K., Croll A., Burnett J.** *Working Life: Employee Attitudes and Engagement 2006*. Chartered Inst. of Personnel and Development, 2006. 54 p.
15. **Fredricks J. A., McColskey W.** The measurement of student engagement: A comparative analysis of various methods and student self-report instruments. *Handbook of research on student engagement/* S. L. Christenson et al. (eds.). Springer Science+Business Media, 2012. Pp. 763–782. doi:10.1007/978-1-4614-2018-7_37
16. **Coates H.** The value of student engagement for higher education quality assurance. *Quality in Higher Education*, 2005, vol. 11, no. 1, pp. 25–36. doi:10.1080/13538320500074915
17. **Greene J. A., Plumley R. D., Urban C. J., Bernacki M. L., Gates K. M., Hogan K. A., Demetriou C., Panter A. T.** Modeling temporal selfregulatory processing in a higher education biology course. *Learning and Instruction*, 2021, vol. 72, pp. 101201. doi:10.1016/j.learninstruc.2019.04.002
18. **Boekaerts M.** Engagement as an inherent aspect of the learning process. *Learning and Instruction*, 2016, vol. 43, pp. 76–83. doi:10.1016/j.learninstruc.2016.02.001
19. **Miller B. W.** Using reading times and eye-movements to measure cognitive engagement. *Educational Psychologist*, 2015, vol. 50, no. 1, pp. 31–42. doi:10.1080/00461520.2015.1004068
20. **Ringeval F., Sonderegger A., Sauer J., Lalanne D.** Introducing the RECOLA multimodal corpus of remote collaborative and affective interactions. *Proc. of IEEE Intern. Conf. and Workshops on Automatic Face and Gesture Recognition (FG)*, 2013, pp. 1–8. doi:10.1109/FG.2013.6553805
21. **Mohamad Nezami O., Dras M., Hamey L., Richards D., Wan S., Paris C.** Automatic recognition of student engagement using deep learning and facial expression. *Joint European Conf. on Machine Learning and Knowledge Discovery in Databases*, 2019, pp. 273–289. doi:10.1007/978-3-030-46133-1_17
22. **Kamath A., Biswas A., Balasubramanian V.** A crowdsourced approach to student engagement recognition in e-learning environments. *Proc. of the 2016 IEEE Winter Conf. on Applications of Computer Vision (WACV)*, 2016, pp. 1–9. doi:10.1109/WACV.2016.7477618
23. **Kaur A., Mustafa A., Mehta L., Dhall A.** Prediction and localization of student engagement in the wild. *Proc. of 2018 Conf. on Digital Image Computing: Techniques and Applications (DICTA)*, 2018, pp. 1–8. doi:10.1109/DICTA.2018.8615851
24. **Gupta A., D’Cunha A., Awasthi K., Balasubramanian V.** DAiSEE: Towards user engagement recognition in the wild. *Journal of Latex Class Files*, 2015, vol. 14, no. 8, 12 p. doi:10.48550/arXiv.1609.01885
25. **Whitehill J., Serpell Z., Lin Y.-Ch., Foster A., Movellan J. R.** The faces of engagement: Automatic recognition of student engagement from facial expressions. *IEEE Transactions on Affective Computing*, 2014, vol. 5, no. 1, pp. 86–98. doi:10.1109/TAFFC.2014.2316163
26. **Celiktutan O., Skordos E., Gunes H.** Multimodal human-human-robot interactions (MHHRI) dataset for studying personality and engagement. *IEEE Transactions on Affective Computing*, 2017, vol. 10, no. 4, pp. 484–497. doi:10.1109/TAFFC.2017.2737019

27. Cafaro A., Wagner J., Baur T., Dermouche S., Torres M. T., Pelachaud C., André E., Valstar M. The NoXi database: multimodal recordings of mediated novice-expert interactions. *Proc. of the 19th ACM Intern. Conf. on Multimodal Interaction*, 2017, pp. 350–359. doi:10.1145/3136755.3136780
28. Guhan P., Agarwal M., Awasthi N., Reeves G., Manocha D., Bera A. ABC-Net: Semi-supervised multimodal GAN-based engagement detection using an affective, behavioral and cognitive model. *arXiv preprint arXiv:2011.08690*, 2020. doi:10.48550/arXiv.2011.08690
29. Psaltis A., Apostolakis K. C., Dimitropoulos K., Daras P. Multimodal student engagement recognition in prosocial games. *IEEE Transactions on Games*, 2018, vol. 10, no. 3, pp. 292–303. doi:10.1109/TCl-AIG.2017.2743341
30. Mayo E. Hawthorne and the western electric company. *The Social Problems of an Industrial Civilization*. Routledge & Kegan Paul, London, 1949. Pp. 161–182.
31. Fedotov D., Perepelkina O., Kazimirova E., Konstantinova M., Minker W. Multimodal approach to engagement and disengagement detection with highly imbalanced in-the-wild data. *Proc. of the Workshop on Modeling Cognitive Processes from Multimodal Data (MCPMD'18)*, 2018, pp. 1–9. doi:10.1145/3279810.3279842
32. Heimerl A., Baur T., André E. A Transparent framework towards the context-sensitive recognition of conversational engagement. *Proc. of the 11th Intern. Workshop on Modelling and Reasoning in Context*, 2020, pp. 7–16.
33. Thomas C., Nair N., Jayagopi D. B. Predicting engagement intensity in the wild using temporal convolutional network. *Proc. of the 20th ACM Intern. Conf. on Multimodal Interaction*, 2018, pp. 604–610. doi:10.1145/3242969.3264984
34. Yang J., Wang K., Peng X., Qiao Y. Deep recurrent multi-instance learning with spatio-temporal features for engagement intensity prediction. *Proc. of the 20th ACM Intern. Conf. on Multimodal Interaction*, 2018, pp. 594–598. doi:10.1145/3242969.3264981
35. Abedi A., Khan S. Affect-driven engagement measurement from videos. *arXiv preprint arXiv:2106.10882*, 2021. doi:10.48550/arXiv.2106.10882
36. Niu X., Han H., Zeng J., Sun X., Shan Sh., Huang Y., Yang S., Chen X. Automatic engagement prediction with GAP feature. *Proc. of the 20th ACM Intern. Conf. on Multimodal Interaction (ICMI'18)*, 2018, pp. 599–603. doi:10.1145/3242969.3264982
37. Dermouche S., Pelachaud C. Engagement modeling in dyadic interaction. *Proc. of the 2019 Intern. Conf. on Multimodal Interaction (ICMI'19)*, 2019, pp. 440–445. doi:10.1145/3340555.3353765
38. Dresvyanskiy D., Minker W., Karpov A. Deep learning based engagement recognition in highly imbalanced data. *Speech and Computer. SPECOM 2021. Lecture Notes in Computer Science*. Springer, Cham, 2021. Vol 12997. Pp. 166–178. doi:10.1007/978-3-030-87802-3_16
39. Szegedy C., Vanhoucke V., Ioffe S., Shlens J., Wojna Z. Rethinking the inception architecture for computer vision. *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, 2016, pp. 2818–2826.
40. O'Brien H., Toms E. The development and evaluation of a survey to measure user engagement. *Journal of the American Society for Information Science and Technology*, 2010, vol. 61, no. 1, pp. 50–69. doi:10.1002/asi.21229
41. Fuller K. A., Karunaratne N. S., Naidu S., Exintaris B., Short J. L., Wolcott M. D., Singleton S., White P. J. Development of a self-report instrument for measuring in-class student engagement reveals that pretending to engage is a significant unrecognized problem. *PLOS ONE*, 2018, vol. 13, no. 10, pp. e0205828. doi:10.1371/journal.pone.0205828
42. Koto I., Harneli M., Winarni E. Primary school teacher strategy to promote student engagement in science lessons. *Proc. of Intern. Conf. of Primary Education Research Pivotal Literature and Research UNNES 2018 (IC PEOPLE UNNES 2018)*, 2018, pp. 122–127. doi:10.2991/icpeopleunnes-18.2019.25
43. Garn A. C., Dasigner T., Simonton A., Simonton K. L. Predicting changes in student engagement in university physical education: Application of control-value theory of achievement emotions. *Psychology of Sport and Exercise*, 2017, no. 29, pp. 93–102. doi:10.1016/j.psychsport.2016.12.005
44. D'Errico F., Paciello M., Cerniglia L. When emotions enhance students' engagement in e-learning processes. *Journal of e-Learning and Knowledge Society*, 2016, vol. 12, no. 4, pp. 173676.
45. Верхоляк О. В., Карпов А. А. Автоматический анализ эмоционально окрашенной речи. *Голосовой портрет ребенка с типичным и атипичным развитием*/ Е. Е. Ляксо, О. В. Фролова (ред.). СПб., Издательско-полиграфическая ассоциация вузов, 2020. С. 149–198.

UDC 004.4/.5+159.99

doi:10.31799/1684-8853-2022-5-12-22

EDN: CXBRCS

Analytical review of methods for automatic detection of user engagement in virtual communicationA. A. Dvoynikova^a, Junior Researcher, orcid.org/0000-0001-8047-6639I. A. Kagirov^a, Research Fellow, orcid.org/0000-0003-1196-1117A. A. Karpov^a, Dr. Sc., Tech., Professor, Principal Researcher, orcid.org/0000-0003-3424-652X, karpov@iias.spb.su^aSt. Petersburg Federal Research Center of the RAS, 39, 14th Line, 199178, Saint-Petersburg, Russian Federation

Introduction: The solution of the task of the recognition and assessment of user engagement in the acts of human-machine interaction or telecommunication, achieved through the use of automatic means, is highly important in computer recognition of human psycho-emotional states. This is necessary for e-learning, business and entertainment applications design. **Purpose:** To conduct a comparative analysis of the current information support in the field of automatic recognition and assessment of user involvement in human-machine interaction or virtual communication, as well as to establish a methodology for building a data body based on the idea of multimodal communication. **Results:** The conducted analysis of research papers has shown that in most existing databases there is a substantial lack of data for natural online communication. Moreover, not all databases contain different modalities in “human – machine – human” communication system. Text and audio modalities turn out to be important for a multilevel engagement classification task, aimed at the determination of engagement intensity. It is also promising to take into account “body language” features, such as facial expressions, movements of the body and the head, gestures. For the correct assessment of involvement, an engagement database must contain meta-data on the psycho-emotional states of communicants. Neural network-based approaches to the automatic detection of user engagement show the best performance. **Practical relevance:** Based on the obtained analytical conclusions, the authors of the paper are going to elaborate an original software system for automatic recognition of user engagement, and to collect a data set for machine learning purposes. The presented review formulates basic requirements for such systems and contributes to the solution of the problem of automatic recognition of psycho-emotional states. **Discussion:** The survey leads to the conclusion that the notion of engagement as understood in studies on automatic emotion recognition differs from that used in psychology. User (or communicant) engagement in terms of info- and communicative sphere implies the manifestation of a person’s mental activity level (emotional, cognitive, and behavioral components) changing dynamically while interacting with another person or computer system.

Keywords – user engagement, information support, automatic emotion recognition systems, multimodality, artificial neural networks.

For citation: Dvoynikova A. A., Kagirov I. A., Karpov A. A. Analytical review of methods for automatic detection of user engagement in virtual communication. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 5, pp. 12–22 (In Russian). doi:10.31799/1684-8853-2022-5-12-22, EDN: CXBRCS

Financial support

This survey was carried out in the framework of the Council for Grants of the President of Russia for Leading scientific schools (grant No. NSH-17.2022.1.6), as well as due as part of Russian state research (No. FFZF-2022-0005).

References

- Pregowska A., Masztalerz K., Garlińska M., Osial M. A worldwide journey through distance education – from the post office to virtual, augmented and mixed realities, and education during the COVID-19 pandemic. *Education Sciences*, 2021, vol. 11, no. 3, pp. 118. doi:10.3390/educsci11030118
- Sümer Ö., Goldberg P., d’Mello S., Gerjets P., Trautwein U., Kasneci E. Multimodal engagement analysis from facial videos in the classroom. *IEEE Transactions on Affective Computing*, 2021, 16 p. doi:10.1109/TAFFC.2021.3127692
- Nkomo L., Daniel B. Sentiment analysis of student engagement with lecture recording. *TechTrends*, 2021, vol. 65, no. 2, pp. 213–224. doi.org:10.1007/s11528-020-00563-8
- Dozortsev V. M., Nazin V. A. Computer simulators as a tool for modeling operator activity in a psychological experiment. *Tr. Instituta psihologii RAN “Aktual’nye problemy psihologii truda, inzhenernoj psihologii i ergonomiki”* [Proc. of the Institute of Psychology of RAS “Actual Problems of Occupational Psychology, Engineering Psychology and Ergonomics”], 2013, iss. 5, pp. 81–103 (In Russian).
- Sokolov V. N., Koroteev G. L. Principles and technologies for building adaptive learning environments. *Tr. Instituta psihologii RAN “Aktual’nye problemy psihologii truda, inzhenernoj psihologii i ergonomiki”* [Proc. of the Institute of Psychology of RAS “Actual Problems of Occupational Psychology, Engineering Psychology and Ergonomics”], 2013, iss. 5, pp. 57–81 (In Russian).
- Dvoynikova A. A., Karpov A. A. Analytical review of approaches to Russian text sentiment recognition *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 4, pp. 20–30 (In Russian). doi:10.31799/1684-8853-2020-4-20-30
- Kahn W. A. Psychological conditions of personal engagement and disengagement at work. *The Academy of Management Journal*, 1990, vol. 33, no. 4, pp. 692–724. doi:10.2307/256287
- Kelders S. M., van Zyl L. E., Ludden G. The concept and components of engagement in different domains applied to eHealth: A systematic scoping review. *Frontiers in Psychology*, 2020, vol. 11, Article 926. doi:10.3389/fpsyg.2020.00926
- De Vreede T., Anel S., de Vreede G.-J., Spector P. E., Singh V., Padmanabhan B. What is engagement and how do we measure it? Toward a domain independent definition and scale. *Proc. of the 52nd Hawaii Intern. Conf. on System Sciences (HICSS 2019)*, 2019, pp. 1–10. doi:10.24251/HICSS.2019.092
- Calder B. J., Malthouse E. C., Schaedel U. An experimental study of the relationship between online engagement and advertising effectiveness. *Journal of Interactive Marketing*, 2009, vol. 23, no. 4, pp. 321–331. doi:10.1016/j.intmar.2009.07.002
- Smith M. An approach to the study of the social act. *Psychological Review*, 1942, vol. 49, no. 5, pp. 422–440. doi:10.1037/h0062907
- Posner M. I. Orienting of attention. *Quarterly Journal of Experimental Psychology*, 1980, vol. 32, no. 1, pp. 3–25. doi:10.1080/00335558008248231
- Li Y., Lerner R. M. Interrelations of behavioral, emotional, and cognitive school engagement in high school students. *Journal of Youth and Adolescence*, 2013, vol. 42, no. 1, pp. 20–32. doi:10.1007/s10964-012-9857-5
- Truss C., Soane E., Edwards C., Wisdom K., Croll A., Burnett J. *Working Life: Employee Attitudes and Engagement 2006*. Chartered Inst. of Personnel and Development, 2006. 54 p.
- Fredricks J. A., McColskey W. *The measurement of student engagement: A comparative analysis of various methods and student self-report instruments*. In: *Handbook of research on student engagement*. S. L. Christenson et al. (eds.). Springer Science+BusinessMedia, 2012. Pp. 763–782. doi:10.1007/978-1-4614-2018-7_37
- Coates H. The value of student engagement for higher education quality assurance. *Quality in Higher Education*, 2005, vol. 11, no. 1, pp. 25–36. doi:10.1080/13538320500074915

17. Greene J. A., Plumley R. D., Urban C. J., Bernacki M. L., Gates K. M., Hogan K. A., Demetriou C., Panter A. T. Modeling temporal selfregulatory processing in a higher education biology course. *Learning and Instruction*, 2021, vol. 72, pp. 101201. doi:10.1016/j.learninstruc.2019.04.002
18. Boekaerts M. Engagement as an inherent aspect of the learning process. *Learning and Instruction*, 2016, vol. 43, pp. 76–83. doi:10.1016/j.learninstruc.2016.02.001
19. Miller B. W. Using reading times and eye-movements to measure cognitive engagement. *Educational Psychologist*, 2015, vol. 50, no. 1, pp. 31–42. doi:10.1080/00461520.2015.1004068
20. Ringeval F., Sonderegger A., Sauer J., Lalanne D. Introducing the RECOLA multimodal corpus of remote collaborative and affective interactions. *Proc. of IEEE Intern. Conf. and Workshops on Automatic Face and Gesture Recognition (FG)*, 2013, pp. 1–8. doi:10.1109/FG.2013.6553805
21. Mohamad Nezami O., Dras M., Hamey L., Richards D., Wan S., Paris C. Automatic recognition of student engagement using deep learning and facial expression. *Joint European Conf. on Machine Learning and Knowledge Discovery in Databases*, 2019, pp. 273–289. doi:10.1007/978-3-030-46133-1_17
22. Kamath A., Biswas A., Balasubramanian V. A crowdsourced approach to student engagement recognition in e-learning environments. *Proc. of the 2016 IEEE Winter Conf. on Applications of Computer Vision (WACV)*, 2016, pp. 1–9. doi:10.1109/WACV.2016.7477618
23. Kaur A., Mustafa A., Mehta L., Dhall A. Prediction and localization of student engagement in the wild. *Proc. of 2018 Conf. on Digital Image Computing: Techniques and Applications (DICTA)*, 2018, pp. 1–8. doi:10.1109/DICTA.2018.8615851
24. Gupta A., D' Cunha A., Awasthi K., Balasubramanian V. DAiSEE: Towards student engagement recognition in the wild. *Journal of Latex Class Files*, 2015, vol. 14, no. 8, 12 p. doi:10.48550/arXiv.1609.01885
25. Whitehill J., Serpell Z., Lin Y.-Ch., Foster A., Movellan J. R. The faces of engagement: Automatic recognition of student engagement from facial expressions. *IEEE Transactions on Affective Computing*, 2014, vol. 5, no. 1, pp. 86–98. doi:10.1109/TAFFC.2014.2316163
26. Celiktutan O., Skordos E., Gunes H. Multimodal human-human-robot interactions (MHHR) dataset for studying personality and engagement. *IEEE Transactions on Affective Computing*, 2017, vol. 10, no. 4, pp. 484–497. doi:10.1109/TAFFC.2017.2737019
27. Cafaro A., Wagner J., Baur T., Dermouche S., Torres M. T., Pelachaud C., André E., Valstar M. The NoXi database: multimodal recordings of mediated novice-expert interactions. *Proc. of the 19th ACM Intern. Conf. on Multimodal Interaction*, 2017, pp. 350–359. doi:10.1145/3136755.3136780
28. Guhan P., Agarwal M., Awasthi N., Reeves G., Manocha D., Bera A. ABC-Net: Semi-supervised multimodal GAN-based engagement detection using an affective, behavioral and cognitive model. *arXiv preprint arXiv:2011.08690*, 2020. doi:10.48550/arXiv.2011.08690
29. Psaltis A., Apostolakis K. C., Dimitropoulos K., Daras P. Multimodal student engagement recognition in prosocial games. *IEEE Transactions on Games*, 2018, vol. 10, no. 3, pp. 292–303. doi:10.1109/TCIAIG.2017.2743341
30. Mayo E. Hawthorne and the western electric company. *The Social Problems of an Industrial Civilization*. Routledge & Kegan Paul, London, 1949. Pp. 161–182.
31. Fedotov D., Perepelkina O., Kazimirova E., Konstantinova M., Minker W. Multimodal approach to engagement and disengagement detection with highly imbalanced in-the-wild data. *Proc. of the Workshop on Modeling Cognitive Processes from Multimodal Data (MCPMD'18)*, 2018, pp. 1–9. doi:10.1145/3279810.3279842
32. Heimerl A., Baur T., André E. A Transparent framework towards the context-sensitive recognition of conversational engagement. *Proc. of the 11th Intern. Workshop on Modelling and Reasoning in Context*, 2020, pp. 7–16.
33. Thomas C., Nair N., Jayagopi D. B. Predicting engagement intensity in the wild using temporal convolutional network. *Proc. of the 20th ACM Intern. Conf. on Multimodal Interaction*, 2018, pp. 604–610. doi:10.1145/3242969.3264984
34. Yang J., Wang K., Peng X., Qiao Y. Deep recurrent multi-instance learning with spatio-temporal features for engagement intensity prediction. *Proc. of the 20th ACM Intern. Conf. on Multimodal Interaction*, 2018, pp. 594–598. doi:10.1145/3242969.3264981
35. Abedi A., Khan S. Affect-driven engagement measurement from videos. *arXiv preprint arXiv:2106.10882*, 2021. doi:10.48550/arXiv.2106.10882
36. Niu X., Han H., Zeng J., Sun X., Shan Sh., Huang Y., Yang S., Chen X. Automatic engagement prediction with GAP feature. *Proc. of the 20th ACM Intern. Conf. on Multimodal Interaction (ICMI'18)*, 2018, pp. 599–603. doi:10.1145/3242969.3264982
37. Dermouche S., Pelachaud C. Engagement modeling in dyadic interaction. *Proc. of the 2019 Intern. Conf. on Multimodal Interaction (ICMI'19)*, 2019, pp. 440–445. doi:10.1145/3340555.3353765
38. Dresvyanskiy D., Minker W., Karpov A. *Deep learning based engagement recognition in highly imbalanced data*. In: *Speech and Computer. SPECOM 2021. Lecture Notes in Computer Science*. Springer, Cham, 2021. Vol 12997. Pp. 166–178. doi:10.1007/978-3-030-87802-3_16
39. Szegedy C., Vanhoucke V., Ioffe S., Shlens J., Wojna Z. Rethinking the inception architecture for computer vision. *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, 2016, pp. 2818–2826.
40. O'Brien H., Toms E. The development and evaluation of a survey to measure user engagement. *Journal of the American Society for Information Science and Technology*, 2010, vol. 61, no. 1, pp. 50–69. doi:10.1002/asi.21229
41. Fuller K. A., Karunaratne N. S., Naidu S., Exintaris B., Short J. L., Wolcott M. D., Singleton S., White P. J. Development of a self-report instrument for measuring in-class student engagement reveals that pretending to engage is a significant unrecognized problem. *PLOS ONE*, 2018, vol. 13, no. 10, pp. e0205828. doi:10.1371/journal.pone.0205828
42. Koto I., Harneli M., Winarni E. Primary school teacher strategy to promote student engagement in science lessons. *Proc. of Intern. Conf. of Primary Education Research Pivotal Literature and Research UNNES 2018 (IC PEOPLE UNNES 2018)*, 2018, pp. 122–127. doi:10.2991/icpeopleunnes-18.2019.25
43. Garn A. C., Dasigner T., Simonton A., Simonton K. L. Predicting changes in student engagement in university physical education: Application of control-value theory of achievement emotions. *Psychology of Sport and Exercise*, 2017, no. 29, pp. 93–102. doi:10.1016/j.psychsport.2016.12.005
44. D'Errico F., Paciello M., Cerniglia L. When emotions enhance students' engagement in e-learning processes. *Journal of e-Learning and Knowledge Society*, 2016, vol. 12, no. 4, pp. 173676.
45. Verkholiyak O. V., Karpov A. A. *Automatic analysis of emotionally charged speech*. In: *Golosovoj portret rebenka s tipichnym i atipichnym razvitiem* [Voice portrait of typical and atypical children]. E. E. Lyakso, O. V. Frolova (Eds.). Saint-Petersburg, Izdatelsko-poligraficheskaya asociaciya vuzov Publ., 2020. Pp. 149–198 (In Russian).



Формирование адаптивных рассылок брокера данных интернета вещей

О. С. Исаева^а, канд. техн. наук, старший научный сотрудник, orcid.org/0000-0002-5061-6765, isaeva@icm.krasn.ru

С. В. Исаев^а, канд. техн. наук, доцент, orcid.org/0000-0002-6678-0084

Н. В. Кулясов^а, программист, orcid.org/0000-0001-5582-9498

^аИнститут вычислительного моделирования СО РАН, Академгородок, 50/44, Красноярск, 660036, РФ

Введение: возможности взаимодействия с физическим миром через сетевые инфраструктуры пространственно-распределенных узлов интернета вещей несмотря на неоспоримые преимущества технологии вызывают существенные нагрузки на потребителей информации. В этой связи актуальным является создание методов, обеспечивающих сокращение передаваемых объемов данных за счет адаптивной синхронизации систем мониторинга со временем протекания реальных процессов. **Цель:** разработать подход к формированию адаптивных рассылок брокера данных на основе исследования цикличности наблюдений устройств интернета вещей. **Результаты:** в рамках корпоративной сети Красноярского научного центра развернута инфраструктура устройств и приложений интернета вещей для мониторинга показателей температуры, влажности и PM2.5 в специализированных технологических помещениях с телекоммуникационным оборудованием. К собираемым данным применен метод дискретного преобразования Фурье. На основании рассчитанных параметров гармонического ряда сделан вывод о частотных характеристиках данных. Выбраны основные пики, описывающие периодичность данных, определены точки колебаний и по теореме Котельникова выбрана частота дискретизации, обеспечивающая достаточную интенсивность наблюдений. Анализ показал, что для различных помещений данные имеют периодический характер, но их гармонические профили не совпадают. Выбор значений гармоник, амплитуда колебания которых определяет динамику изменений в наблюдаемых данных, следует проводить периодически для каждого наблюдаемого устройства. Этот подход реализован в программном обеспечении брокера, который выдает данные по подпискам от каждого из устройств в соответствии с частотой их изменений. **Практическая значимость:** анализ частотных характеристик данных определяет настройки брокера, которые сокращают потоки выдаваемой информации, что является одним из аспектов обеспечения надежности инфраструктуры интернета вещей. Кроме того, наблюдение за изменениями характера данных позволяет выявлять неполадки в работе охлаждающих систем, которые могут привести к выходу из строя сложного дорогостоящего оборудования, обладающего повышенной теплоотдачей.

Ключевые слова — интернет вещей, протокол обмена сообщениями, Message Queuing Telemetry Transport, дискретное преобразование Фурье, теорема Котельникова.

Для цитирования: Исаева О. С., Исаев С. В., Кулясов Н. В. Формирование адаптивных рассылок брокера данных интернета вещей. *Информационно-управляющие системы*, 2022, № 5, с. 23–31. doi:10.31799/1684-8853-2022-5-23-31, EDN: DNOSCW

For citation: Isaeva O. S., Isaev S. V., Kulyasov N. V. Formation of adaptive publications from the Internet of Things data broker. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 5, pp. 23–31 (In Russian). doi:10.31799/1684-8853-2022-5-23-31, EDN: DNOSCW

Введение

Концепция интернета вещей (Internet of Things, IoT) как обобщающего ключевого понятия для «вещь-ориентированных», «интернет-ориентированных» и «семантико-ориентированных» технологий обеспечивает расширение базовых основ Интернета в физическую сферу [1, 2]. Возможность взаимодействия с физическим миром осуществляется благодаря наличию устройств, способных ощущать физические явления и преобразовывать их в поток информации, а также способных запускать действия, влияющие на физические устройства через исполнительные механизмы. Таким образом, IoT представляет собой динамическую глобальную сетевую инфраструктуру пространственно-распределенных узлов, оснащенных встроенными средствами для взаимодействия друг с другом

или с внешней средой и способных измерять, понимать и даже изменять свое окружение [3].

По разным оценкам международных аналитических агентств, количество активных устройств IoT в ближайшее время в мире достигнет от 30 до 60 млрд [4]. Такие цифры предполагают, что IoT станет одним из основных источников больших данных [5], что влечет за собой необходимость переосмысления традиционных подходов к организации вычислительных ресурсов и услуг. Решения проблем надежности, производительности, безопасности и конфиденциальности взаимодействия всех уровней сетевой архитектуры поставщиков или потребителей информации лежат на стыке таких областей знаний, как телекоммуникации, информатика, электроника и др. [6]. И такое решение основывается на детальном исследовании архитектуры IoT и протоколов передачи данных.

В общем случае архитектура IoT состоит из следующих функциональных уровней: сенсорного, транспортного, сервисного и прикладного [7]. Уровни и объекты IoT обладают свойством неоднородности и могут содержать дополнительные приложения для обслуживания, управления и использования устройств, включая широкий спектр интеллектуальных систем. Благодаря такой неоднородности IoT становится важным аспектом повседневной жизни [8, 9]. На сенсорном уровне размещаются датчики, обеспечивающие сбор информации о состоянии наблюдаемых объектов. Устройства IoT имеют ограниченные ресурсы, которые не допускают сложной обработки данных на месте. Для их передачи разворачивается транспортная структура, включающая шлюзы и сети передачи данных. Через этот уровень от устройств IoT поступают данные. Они интегрируются на сервисном уровне, где размещается брокер данных, обеспечивающий сбор, интеграцию, хранение и передачу информации в соответствии с используемыми протоколами. На этом уровне консолидируются неструктурированные или полуструктурированные данные, которые характеризуются большим объемом, разнообразием и частотой генерации [10]. Подготовленные данные поступают на прикладной уровень, включающий проблемно-ориентированные приложения, решающие задачи конкретной предметной области [11]. Для обмена данными используются различные протоколы. В нашем исследовании выбран открытый протокол MQTT (Message Queue Telemetry Transport) [12]. Он позволяет работать с удаленными локациями, имеющими ограничения по возможной обработке и пропускной способности каналов связи. Реализация протокола предусматривает взаимодействие логических сущностей со следующими ролями: Издатель (Publisher) – устройство IoT, формирующее сообщение; Брокер (Broker) – специализированное программное обеспечение, получающее и распределяющее сообщения, например Mosquitto MQTT, и Подписчик (Subscriber) – устройства или программное обеспечение, получающее данные от разных издателей по заданным тематическим подпискам [13, 14].

Гетерогенность и иерархичность архитектуры IoT обуславливают необходимость создания специализированных подходов к анализу безопасности, а также к организации механизмов сбора и мониторинга генерируемых данных. В работах [15, 16] показаны подходы к мониторингу и анализу сетевого трафика структурированных данных журналов обращений. В [17, 18] проводится анализ уязвимостей, определяемых особенностями аутентификации в протоколе MQTT схемы Издатель – Брокер – Подписчик, однако полу-

ченные в них рекомендации носят общий характер и не предоставляют автоматизированных подходов. В [19, 20] выполнено исследование сетевой архитектуры IoT и предложены универсальные инструменты анализа безопасности различных уровней сети IoT. В их основе лежит подход к построению профилей умных устройств из статистических характеристик сеансов связи: интенсивности и продолжительности передачи пакетов данных. Предложен единый (для всех уровней архитектуры IoT) подход к обнаружению атак, основанный на методах машинного обучения. Проблемы адаптивного доступа к данным на прикладном уровне архитектуры IoT, как правило, остаются за рамками проводимых исследований. Существенное значение эта проблема имеет при использовании конечных устройств, обладающих ограниченными возможностями по обработке данных (мобильных сервисов и приложений). Разумное сокращение передаваемых объемов информации является одним из аспектов защиты от перегрузки сети IoT и способствует обеспечению ее надежности и безопасности. В этом случае требуется автоматизировать анализ поступающих от устройств IoT пакетов, определять скорости протекания наблюдаемых событий и выполнять адаптивную настройку режима выдачи информации от брокера данных потребителям.

Цель исследования – разработать подход к формированию адаптивных рассылок брокера данных на основе исследования данных, собираемых устройствами IoT.

Поскольку большинство процессов как для производственных, так и для бытовых систем имеют периодический характер, то изучение циклических характеристик данных, собираемых устройствами IoT, оправдано и позволит лучше понимать суть происходящих в них явлений. Цикличность – это свойство работы системы, которая стремится сохранить свое состояние в границах равновесия. Для исследования в работе применен метод дискретного преобразования Фурье [21], и на основе рассчитанных параметров гармонического ряда сделан вывод о частотных характеристиках данных. Подобный подход получил свое применение в работах по исследованию частотно-временных характеристик различных последовательностей данных [22].

Постановка задачи исследования данных IoT

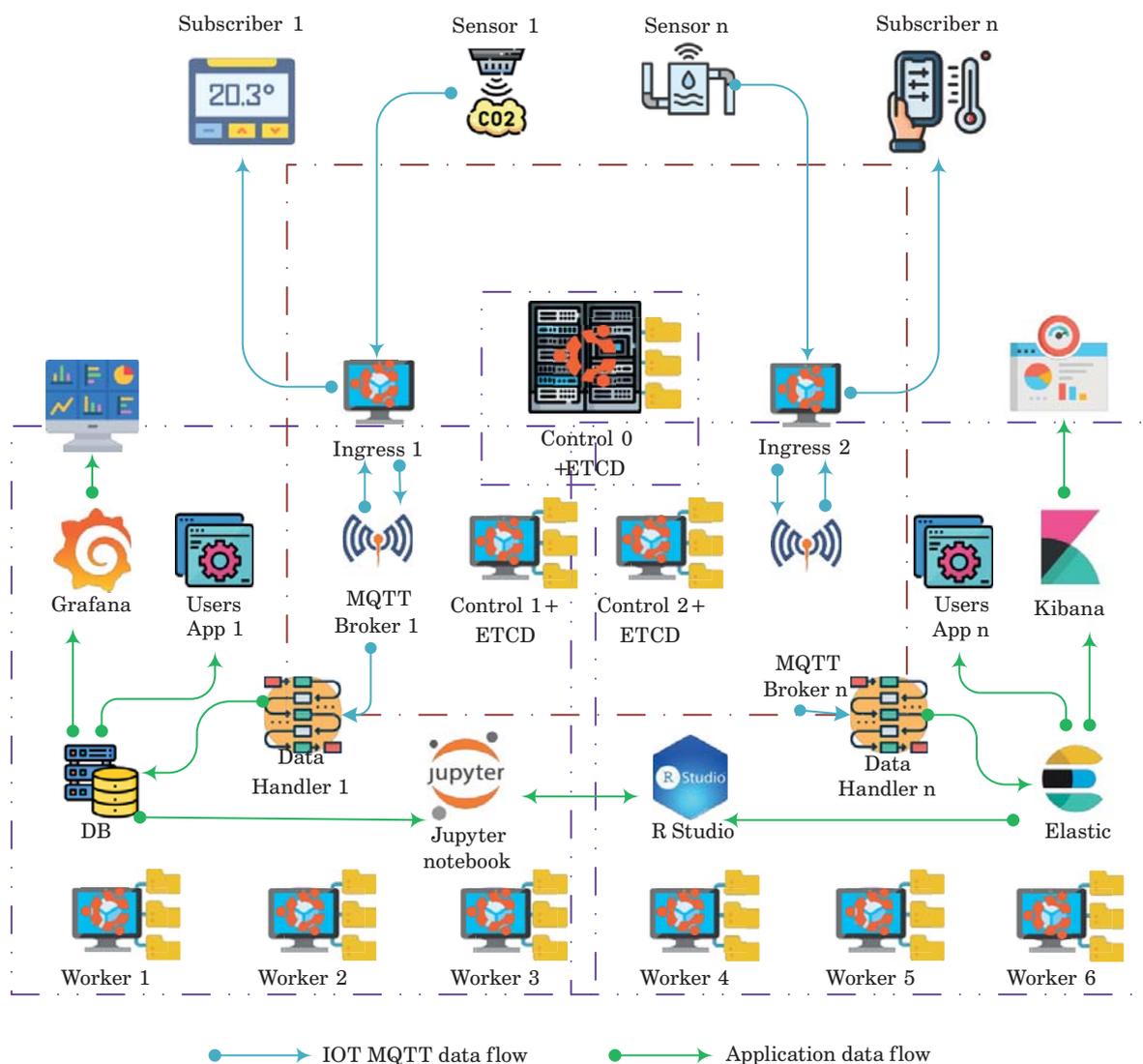
Для проведения исследований в рамках корпоративной сети Красноярского научного центра развернута инфраструктура устройств и приложений IoT. Устройства IoT размещены в специализированных технологических помещениях с теле-

коммуникационным оборудованием и выполняют мониторинг показателей температуры, влажности и PM2.5. Результаты измерений датчиков от устройств IoT по протоколу MQTT поступают брокерам, которые их собирают, обрабатывают, размещают в базах данных и передают в устройства и приложения. Кроме того, брокер собирает сетевые журналы, содержащие сведения о наличии обращений к данным и несанкционированных запросах. У подписчиков (например, мобильных приложений) существуют ограничения на объемы и частоту поступления данных. Задача брокера – обеспечить адаптивность режима выдачи данных, при котором частота их обновления соответствовала бы скорости протекания наблюдаемых событий. Требуется выполнить анализ поступающих данных и создать инструменты для формирования адаптивных рассылок.

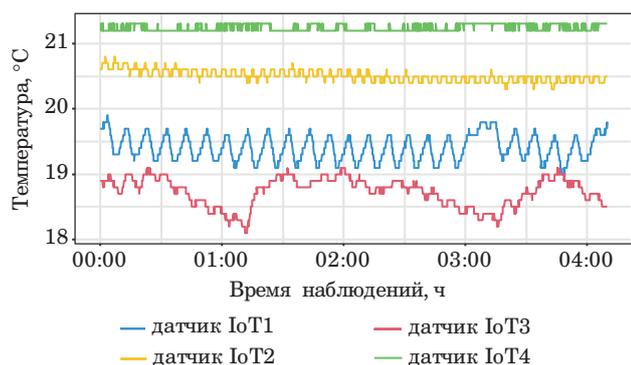
Концептуальная схема инфраструктуры IoT (рис. 1) содержит как развернутые в настоящий момент системы и сервисы, так и потенциальные точки расширения и объединяет потоки данных устройств IoT и приложений.

В концептуальной схеме представлены основные устройства и приложения инфраструктуры:

- устройства сенсорного уровня для сбора данных и взаимодействия с окружением по протоколу MQTT (на рисунке обозначены как Sensor);
- маршрутизаторы трафика из внешней сети в кластер (Ingress);
- узлы кластера, предназначенные для развертывания контейнеров с приложениями (Control, Worker);
- брокеры рассылки данных по протоколу MQTT (MQTT Broker);



■ **Рис. 1.** Концептуальная схема инфраструктуры IoT
 ■ **Fig. 1.** Conceptual diagram of the IoT infrastructure



■ **Рис. 2.** График температуры с датчиков
 ■ **Fig. 2.** Graph of temperature from sensors

- сервисы для анализа данных и настройки рассылок (Data Handler);
- база данных (DB – в нашей реализации MySQL и PhpMyAdmin);
- обработчики данных и приложения, выступающие подписчиками MQTT-брокера (Subscriber);
- средства разработки, графические интерфейсы и компоненты статистического анализа и визуализации (R-Studio, Jupiter, Kibana, Grafana, Elastik, Users App).

Детали организации и настройки самой инфраструктуры IoT в настоящей статье не рассматриваются.

Из приведенной структуры выделены потоки данных между издателями, брокерами и подписчиками. Устройства IoT собирают данные с датчиков (температуры, влажности и PM2.5), на рис. 2 показан фрагмент графика результатов измерений, получаемых брокером данных от устройств IoT.

Из наблюдений видно, что характеристики цикличности результатов измерений различаются для каждого из устройств.

Формирование адаптивных рассылок

Устройства IoT установлены в технологических помещениях с оборудованием, обладающим повышенной теплоотдачей. В помещениях установлены системы кондиционирования, выход из строя или изменение режима работы которых повлечет проблемы в работоспособности сложного коммуникационного и вычислительного оборудования. Периодичность поступления данных с устройств IoT в MQTT-брокер настраивается на источниках данных.

Рассмотрим дискретные моменты времени наблюдения $\{t_n\}$, где $n = [0, N - 1]$, N – количество отсчетов. Время между наблюдениями

$\Delta t = (t_{n+1} - t_n) + \tau$, где τ – задержки передачи данных. Значение τ столь мало, что период дискретизации можно считать постоянным, равным $T = (t_{n+1} - t_n)$. Длительность наблюдений равна NT . Представим последовательность наблюдаемых данных как функцию дискретного аргумента, принимающую произвольные положительные значения в дискретные промежутки времени nT : $x(t) = \{x(nT)\}$, где $x(nT)$ – n -й результат наблюдений в момент nT .

Функция $x(t)$ – периодическая, в общем случае свойство периодичности достигается путем повторения рассматриваемых данных с периодом NT . Для периодической функции может быть выполнено спектральное разложение в виде ряда Фурье и получен дискретный спектр $X(k)$, $k \in Z$, состоящий из гармоник, кратных $\Delta\omega = 2\pi/T$. Поскольку исходная функция $x(t)$ – дискретная, достаточно определить спектральные значения для $k = [0, N - 1]$. Коэффициенты разложения в ряд Фурье могут быть получены по формуле дискретного преобразования, которое зависит только от индекса входного сигнала:

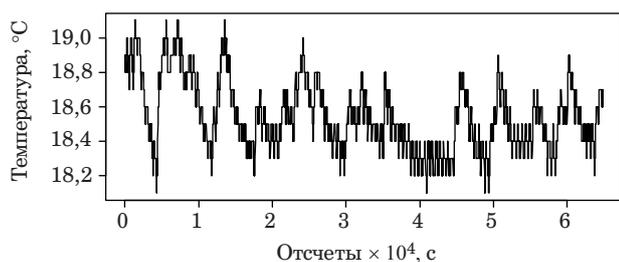
$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-i \frac{2\pi}{N} nk} \quad (1)$$

Преобразование позволяет по N измерениям значений $x(t)$ получить N спектральных отсчетов на одном периоде повторения спектра $X(k)$.

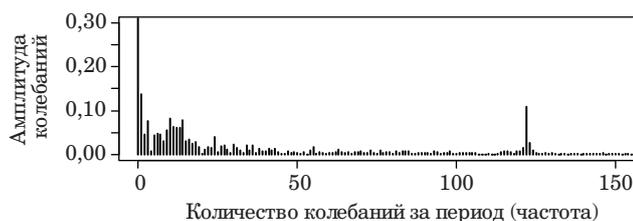
Применим к нашим результатам наблюдений по каждому из источников дискретное преобразование Фурье, рассмотрим модули комплексных чисел полученных коэффициентов. По ним будем делать вывод о значениях амплитуд и частот гармоник. На основе полученного результата сделаем вывод о частоте изменения наблюдаемых событий и сформируем настройки для выдачи данных, которые позволят выбирать значимые события. Такую последовательность действий требуется выполнять для большого набора данных с заданной периодичностью, чтобы оперативно изменять настройки времени событий выдачи данных. Реализация этого подхода позволит автоматически формировать адаптивные подписки.

Поясним подход к формированию адаптивных рассылок на примере. Возьмем выборку данных, являющихся результатом измерений в N отсчетах (рис. 3).

Применим дискретное преобразование Фурье (1) и построим частотный спектр данных. Весовые коэффициенты разложения являются комплексным спектром периодического сигнала (нашей выборки данных). По коэффициентам Фурье построим $\text{mod}(X(k))$, где mod – функция, возвращающая действительную часть комплексного числа. График результатов показан на рис. 4.



■ **Рис. 3.** Данные в наблюдаемых отсчетах
 ■ **Fig. 3.** Data in the observed samples



■ **Рис. 4.** Частотный спектр данных
 ■ **Fig. 4.** Frequency spectrum of data

Построенный спектр — дискретный и состоит из гармоник, кратных $\Delta\omega$. Его первая гармоника при $k = 0$ является основной частотой сигнала и отражает постоянную составляющую данных. Для настройки адаптивных рассылок ее не учитываем. Остальные частоты дискретного спектра (при $k \geq 1$) являются гармониками сигнала. Требуется найти значения гармоник, амплитуда колебания которых показывает динамику изменений в наблюдаемых данных. Зададим порог ε для выбора значений амплитуд, которые будем учитывать при анализе. Минимальное ε может быть определено из характеристик точности используемых датчиков IoT. Выберем коэффициенты разложения $X(k)$ такие, что $\text{Re}(X(k)) > \varepsilon$ и выполнено условие

$$X(k - 1) \leq X(k) \geq X(k + 1). \quad (2)$$

По k (номеру коэффициента) определим период колебаний, соответствующий выбранным $X(k)$, разделив интервал наблюдений NT на номер отсчета:

$$P(k) = NT/k. \quad (3)$$

Частоту колебаний будем рассчитывать по формуле

$$F(k) = 1/P(k). \quad (4)$$

Выберем для k , удовлетворяющих условию (2), максимальную частоту колебаний спектра сигнала, определяющую период дискретизации Fd :

$$Fd = \max F(k). \quad (5)$$

Воспользуемся теоремой Котельникова [23, 24], согласно которой если спектр сигнала ограничен частотой F_{\max} , то он может быть однозначно восстановлен по его дискретным отсчетам, взятым через интервалы времени с частотой как минимум в два раза превышающей максимальную частоту сигнала, который мы хотим измерить, т. е. на каждое колебание сигнала (изменение измеренных значений) должно приходиться как минимум два отсчета. Отсюда следует, что для наблюдения за данными достаточно выполнять измерения с частотой дискретизации

$$F_{\max} \geq 2Fd. \quad (6)$$

Минимальный период дискретизации определяется обратно максимальной частоте:

$$P_{\min} = 1/F_{\max}. \quad (7)$$

В рассматриваемом примере значения $k \in \{10, 14, 122\}$ удовлетворяют условию (2) коэффициентов разложения (1) с частотами, соответствующими периодам колебаний (3) $P(k) \in \{6480, 4632, 525\}$ секунд соответственно. По (4) определяем частоту. По (5) выбираем максимальную частоту колебаний наблюдаемых данных: $Fd = 1/525$ Гц, и частоту дискретизации можем выбрать из (6): $F_{\max} = Fd \cdot 2 = 2/525$ Гц. Отсюда получаем по (7) период дискретизации $P_{\min} = 262,5$ с. Найденный период P_{\min} применяется для настройки брокера, что обеспечивает наблюдение за событиями в соответствии с динамикой их изменений.

Выбор периода дискретизации данных для устройств IoT

Предложенный подход применим для настройки выдачи данных от устройств IoT, в том числе для энергонезависимых источников. Уменьшение частоты пересылки данных от таких устройств к брокеру и увеличение длительности периодов покоя продлит время их автономного функционирования.

Выбор параметров дискретизации измерений выполняется на основе статистики их работы по формуле (7). Рассматривается длительный интервал наблюдений. Для избежания потери информативности в моменты неактивности устройств IoT при выборе интервала дискретизации следует учитывать минимальный период измерений при максимальной скорости протекания наблюдаемых процессов. В общем случае скорость может быть получена на основе собран-



■ **Рис. 5.** Данные для расчета максимальной скорости протекания процессов

■ **Fig. 5.** Data for calculating the maximum speed of processes

ной статистики (при имитации критических состояний):

$$S_{\max} = \max_{k=1, N} \frac{|X(k) - X(k-1)|}{\Delta t}. \quad (8)$$

Для каждого наблюдаемого показателя зададим значение X_{\min} его допустимого изменения, тогда период вычисляется на основе скорости протекания процессов:

$$P_S = X_{\min} / S_{\max}. \quad (9)$$

Результирующий период дискретизации для работы устройств IoT определяется из результатов (7) и (9):

$$P_d = \min(P_S, P_{\min}). \quad (10)$$

Рассмотрим пример определения периода дискретизации для датчиков IoT измерения температуры окружающей среды. Устройства IoT выполняют измерения и передают данные брокеру каждые 10 с. Требуется поддерживать температуру технологических помещений в диапазоне 18–24 °C. Покажем, каким должен быть режим выдачи данных при допустимом изменении на 2 °C.

Скорость протекания процессов вычисляется из данных, полученных как при стандартном режиме работы, так и при имитации выхода из строя систем охлаждения и увеличении нагрузки на серверное оборудование. Фрагмент графика исходных данных приведен на рис. 5.

По (8) максимальная скорость изменения температуры составляет 0,04 °C/с, по (9), (10) период дискретизации составляет 50 с. Проведенный расчет позволяет сократить количество передаваемых данных и периодов активности устройств IoT в 5 раз. Перенастройка устройств IoT должна происходить при изменении параметров физического состояния окружения.

Обсуждение предложенного подхода

Реализация предложенного в работе подхода выполнена в виде сервиса на языке R. Используемый брокер имеет открытый код и модифицирован для получения конфигурационных настроек от сервиса. Получатели данных подписываются на данные брокера без возможности взаимодействовать напрямую с устройствами IoT.

Подход применим для стандартного функционирования, когда необходимо снизить нагрузку на мобильные устройства, выполняющие мониторинг данных. Он позволяет сократить объем рассылки от брокера и уменьшить загрузку каналов связи при сохранении адекватности представления протекающих процессов. Такие рассылки помечаются флагом «0 – at most once» в поле качества обслуживания Quality of Service (QoS) протокола MQTT. Сообщения, выходящие за границы диапазонов безопасного функционирования систем, не фильтруются данным сервисом и помечаются флагом «2 – exactly once», что гарантирует их обязательную доставку подписчикам. В дальнейшем сервисы обработки данных будут модифицированы для обнаружения критических ситуаций и отправки сообщений брокером с измененными флагами QoS.

В настоящий момент рассматриваются варианты внедрения функций сервиса в код брокера. Это возможно, поскольку в расчетах используется быстрое дискретное преобразование Фурье, имеющее сложность $O(N \cdot \log N)$, что не приводит к существенным вычислительным затратам, в том числе и при увеличении длительности периода, за который анализируются данные (подтверждено экспериментами). В рассмотренном примере длительность периода выборки данных, превышающая сутки, не актуальна, так как протекающие процессы имеют периодичность несколько десятков секунд. Но этот подход может быть применен для медленно протекающих процессов, где период рассылки может быть существенно увеличен, что потребует увеличения периода выборки анализируемых данных.

При наблюдении нескольких параметров следует выбирать для них минимальный период дискретизации. Например, для показателей влажности технологических помещений существенным критерием является скорость изменения показателя не более 6 % в час для исключения конденсации влаги. Эта оценка устанавливается для расчета периода дискретизации для устройств измерения влажности. К непериодическим процессам или процессам, имеющим мгновенное протекание, предложенный в работе подход не применим.

Заключение

Интернет вещей определяет новые архитектурные и технологические решения для разнообразных инфокоммуникационных задач. Выполненное исследование по формированию адаптивных рассылок данных устройств IoT является основой функционирования автономных систем сбора данных и независимых сервисов для их анализа. Применение математических методов сокращает объемы передаваемых данных и оптимизирует работу всех элементов инфраструктуры IoT, каждое устройство которой имеет собственные спектральные характеристики. Автоматизация указанных функций позволяет соотносить время наблюдений со временем протекания реальных процессов. Анализ частотных характеристик собранных данных выполняется в заданные промежутки времени и определяет адаптивность персональных настроек брокера данных по всем анализируемым устройствам, что позволяет сократить нагрузку на устройства – потребители информации.

Дальнейшее исследование спектральных характеристик собираемых данных позволит понимать суть наблюдаемых явлений и выявлять нарушения в работе устройств, оказывающих

влияние на измеряемые показатели. Вследствие этого применение данного подхода для мониторинга состояния специализированных технологических помещений с телекоммуникационным оборудованием имеет существенное практическое значение. Так, например, изменение спектральных характеристик температурных наблюдений даже в случае нахождения всех значений в допустимых границах свидетельствует о необходимости провести диагностику охлаждающего оборудования, поскольку может быть следствием отключения части устройств или изменения их режима работы. Превентивные меры позволяют избежать неконтролируемого превышения температурных значений, которое приведет к выходу из строя сложного дорогостоящего оборудования, обладающего повышенной теплоотдачей.

Финансовая поддержка

Работа поддержана Красноярским математическим центром, финансируемым Минобрнауки РФ в рамках мероприятий по созданию и развитию региональных НОМЦ (соглашение № 075-02-2022-873).

Литература

1. Atzori L., Iera A., Morabito G. The Internet of Things: A survey. *Computer Networks*, 2010, no. 54, pp. 2787–2805. doi:10.1016/j.comnet.2010.05.010
2. Korte A., Tiberius V., Potsdam U., Brem A., Stuttgart U. Internet of Things (IoT) technology research in business and management literature: Results from a co-citation analysis. *Journal of Theoretical and Applied Electronic Commerce Research*, 2021, no. 16(6), pp. 2073–2090. doi:10.3390/jtaer16060116
3. Rozik A. S., Tolba A. S., El-Dosuky M. A. Design and implementation of the Sense Egypt platform for real-time analysis of IoT data streams. *Advances in Internet of Things*, 2016, no. 6(4), pp. 66–91.
4. Alam T. A reliable communication framework and its use in Internet of Things (IoT). *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 2018, no. 3(5), pp. 450–456.
5. Gubbi J., Buyya R., Marusic S., Palaniswami M. Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems*, 2013, no. 29(7), pp. 1645–1660. doi:10.1016/j.future.2013.01.010
6. D'Angelo G., Ferretti S., Ghini V. Simulation of the Internet of Things. *Proc. of the Intern. Conf. on High Performance Computing and Simulation*, 2016, pp. 1–8. doi:10.1109/HPCSim.2016.7568309
7. Javed A., Heljanko K., Buda A., Främling K. CEFIoT: A fault-tolerant IoT architecture for edge and cloud. *IEEE World Forum on Internet of Things*, 2018, pp. 813–818. doi:10.1109/WF-IoT.2018.8355149
8. Botta A., Donato W., Persico V., Pescapé A. Integration of cloud computing and Internet of Things: A survey. *Future Generation Computer Systems*, 2016, no. 56, pp. 684–700. doi:10.1016/j.future.2015.09.021
9. Kumar S., Tiwari P., Zymbler M. Internet of Things is a revolutionary approach for future technology enhancement: A review. *Journal of Big Data*, 2019, no. 6, pp. 1–21. doi:10.1186/s40537-019-0268-2
10. Azad P., Navimipour N. J., Rahmani A. M. The role of structured and unstructured data managing mechanisms in the Internet of Things. *Cluster Computing*, 2020, no. 23, pp. 1185–1198. doi:10.1007/s10586-019-02986-2
11. Sanabria-Russo L., Pubill D., Serra J., Verikoukis C. IoT data analytics as a network edge service. *IEEE Conf. on Computer Communications Workshops*, 2019, pp. 969–970. doi:10.1109/INFCOMW.2019.8845207
12. Dizdarević J., Carpio F., Jukan A., Masip-Bruin X. A survey of communication protocols for Internet of Things and related challenges of fog and cloud computing integration. *ACM Computing Surveys*, 2019, vol. 51, iss. 6, pp. 1–29.
13. *Mosquitto-MQTT-broker*. <https://github.com/topics/mosquitto-mqtt-broker> (дата обращения: 11.07.2022).

14. Patel C., Doshi N. A novel MQTT security framework in generic IoT model. *Procedia Computer Science*, 2020, no. 171, pp. 1399–1408. doi:10.1016/j.procs.2020.04.150
15. Isaev S., Kononov D. Analysis of the dynamics of Internet threats for corporate network web services. *CEUR Workshop Proceedings*, 2021, no. 3047, pp. 71–78. doi:10.47813/sibdata-2-2021-10
16. Kononov D. D., Isaev S. V. Development of secure automated management systems based on web technologies. *IOP Conference Series: Materials Science and Engineering*, 2019, no. 537(5). doi:10.1088/1757-899X/537/5/052024
17. Дикий Д. И., Артемьева В. Д. Протокол передачи данных MQTT в модели удаленного управления правами доступа для сетей Интернета. *Научно-технический вестник информационных технологий, механики и оптики*, 2019, т. 19, № 1, с. 109–117. doi:10.17586/2226-1494-2019-19-1-109-117
18. Andy S., Rahardjo B., Hanindhito B. Attack scenarios and security analysis of MQTT communication protocol in IoT system. *4th Intern. Conf. on Electrical Engineering, Computer Science and Informatics*, 2017, pp. 1–6. doi:10.1109/EECSI.2017.8239179
19. Татарникова Т. М., Богданов П. Ю. Обнаружение атак в сетях интернета вещей методами машинного обучения. *Информационно-управляющие системы*, 2021, № 6, с. 42–52. doi:10.31799/1684-8853-2021-6-42-52
20. Meidan Y., Bohadana M., Mathov Y., Mirsky Y., Breitenbacher D., Shabtai A., Elovici Y. N. BaIoT: Network-based detection of IoT botnet attacks using deep autoencoders. *IEEE Pervasive Computing, Special Issue – Securing the IoT*, 2018, no. 17(3), pp. 12–22.
21. Duhamel P., Vetterli M. Fast Fourier transforms: a tutorial review and a state of the art. *Signal Processing*, 1990, no. 19, pp. 259–299.
22. Кононович Э. В., Миронова И. В., Батулин В. А. Частотно-временной анализ рядов солнечной активности. *Исследовано в России*, 2006, № 182, с. 1704–1715. <http://crydee.sai.msu.ru/~mir/182.pdf> (дата обращения: 13.05.2022).
23. Зиятдинов С. И. Восстановление сигнала по его выборкам на основе теоремы отсчетов Котельникова. *Изв. вузов. Приборостроение*, 2010, № 53(5), с. 44–47.
24. Lüke H. D. The origins of the sampling theorem. *IEEE Communications Magazine*, 1999, no. 37(4), pp. 106–108.

UDC 004.6

doi:10.31799/1684-8853-2022-5-23-31

EDN: DNOSCW

Formation of adaptive publications from the Internet of Things data brokerO. S. Isaeva^a, PhD, Tech., Senior Researcher, orcid.org/0000-0002-5061-6765, isaeva@icm.krasn.ruS. V. Isaev^a, PhD, Tech., Associate Professor, orcid.org/0000-0002-6678-0084N. V. Kulyasov^a, Programmer, orcid.org/0000-0001-5582-9498^aInstitute of Computational Modelling SB RAS, 50/44, Akademgorodok St., 660036, Krasnoyarsk, Russian Federation

Introduction: The possibility of interaction with the physical world through the network infrastructures of spatially distributed nodes of Internet of Things (IoT), despite the undeniable advantages of the technology, produces significant loads on information consumers. In this regard, the current interest is the creation of methods that provide the reduction of transmitted data due to the adaptive synchronization of monitoring systems with the time of real processes. **Purpose:** To develop an approach to the formation of adaptive data broker publications based on the study of the cyclicity of observations of Internet of Things devices. **Results:** Within the corporate network of Krasnoyarsk Research Center, an infrastructure of devices and applications of the Internet of Things has been deployed to monitor temperature, humidity and PM2.5 in specialized technological rooms with telecommunications equipment. The discrete Fourier transform method was applied to the collected data. Based on the calculated parameters of the harmonic series, a conclusion has been made about the frequency characteristics of the data. The main peaks describing the periodicity of the data have been selected, the oscillation points have been determined, and, according to the Nyquist – Shannon – Kotelnikov theorem, a sampling frequency that provides a sufficient intensity of observations has been chosen. The analysis has shown that for different rooms the data are periodic but their harmonic profiles do not coincide. The choice of harmonic values whose oscillation amplitude determines the dynamics of changes in the data observed should be carried out periodically for each device under observation. This approach is implemented in the broker software which distributes data in subscriptions from each of the devices in accordance with the frequency of their changes. **Practical relevance:** The analysis of the frequency characteristics of data determines the broker settings that reduce output information flows, which is one of the aspects of ensuring the reliability of the IoT infrastructure. In addition, observing data changes allows to detect cooling system operation faults and malfunctions which can lead to the failure of sophisticated and expensive equipment with increased heat irradiation.

Keywords – Internet of Things, Message Queuing Telemetry Transport, discrete Fourier transform, Nyquist – Shannon – Kotelnikov sampling theorem.

For citation: Isaeva O. S., Isaev S. V., Kulyasov N. V. Formation of adaptive publications from the Internet of Things data broker. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 5, pp. 23–31 (In Russian). doi:10.31799/1684-8853-2022-5-23-31, EDN: DNOSCW

Financial support

This work is supported by the Krasnoyarsk Mathematical Center and financed by the Ministry of Science and Higher Education of the Russian Federation in the framework of the establishment and development of regional Centers for Mathematics Research and Education (Agreement No. 075-02-2022-873).

References

- Atzori L., Iera A., Morabito G. The Internet of Things: A survey. *Computer Networks*, 2010, no. 54, pp. 2787–2805. doi:10.1016/j.comnet.2010.05.010
- Korte A., Tiberius V., Potsdam U., Brem A., Stuttgart U. Internet of Things (IoT) technology research in business and management literature: Results from a co-citation analysis. *Journal of Theoretical and Applied Electronic Commerce Research*, 2021, no. 16(6), pp. 2073–2090. doi:10.3390/jtaer16060116
- Rozik A. S., Tolba A. S., El-Dosuky M. A. Design and implementation of the Sense Egypt platform for real-time analysis of IoT data streams. *Advances in Internet of Things*, 2016, no. 6(4), pp. 66–91.
- Alam T. A reliable communication framework and its use in Internet of Things (IoT). *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 2018, no. 3(5), pp. 450–456.
- Gubbi J., Buyya R., Marusic S., Palaniswami M. Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems*, 2013, no. 29(7), pp. 1645–1660. doi:10.1016/j.future.2013.01.010
- D'Angelo G., Ferretti S., Ghini V. Simulation of the Internet of Things. *Proc. of the Intern. Conf. on High Performance Computing and Simulation*, 2016, pp. 1–8. doi:10.1109/HPC-Sim.2016.7568309
- Javed A., Heljanko K., Buda A., Främling K. CEFIoT: A fault-tolerant IoT architecture for edge and cloud. *IEEE World Forum on Internet of Things*, 2018. pp. 813–818. doi:10.1109/WF-IoT.2018.8355149
- Botta A., Donato W., Persico V., Pescapé A. Integration of cloud computing and internet of things: A survey. *Future Generation Computer Systems*, 2016, no. 56, pp. 684–700. doi:10.1016/j.future.2015.09.021
- Kumar S., Tiwari P., Zymbler M. Internet of Things is a revolutionary approach for future technology enhancement: A review. *Journal of Big Data*, 2019, no. 6, pp. 1–21. doi:10.1186/s40537-019-0268-2
- Azad P., Navimipour N. J., Rahmani A. M. The role of structured and unstructured data managing mechanisms in the Internet of Things. *Cluster Computing*, 2020, no. 23, pp. 1185–1198. doi:10.1007/s10586-019-02986-2
- Sanabria-Russo L., Pubill D., Serra J., Verikoukis C. IoT data analytics as a network edge service. *IEEE Conf. on Computer Communications Workshops*, 2019, pp. 969–970. doi:10.1109/INFCOMW.2019.8845207
- Dizdarević J., Carpio F., Jukan A., Masip-Bruin X. A survey of communication protocols for Internet of Things and related challenges of fog and cloud computing integration. *ACM Computing Surveys*, 2019, vol. 51, iss. 6, pp. 1–29.
- Mosquito-MQTT-broker. Available at: <https://github.com/topics/mosquito-mqtt-broker> (accessed 11 July 2022).
- Patel C., Doshi N. A novel MQTT security framework in generic IoT model. *Procedia Computer Science*, 2020, no. 171, pp. 1399–1408. doi:10.1016/j.procs.2020.04.150
- Isaev S., Kononov D. Analysis of the dynamics of Internet threats for corporate network web services. *CEUR Workshop Proceedings*, 2021, no. 3047, pp. 71–78. doi:10.47813/sibdata-2-2021-10
- Kononov D. D., Isaev S. V. Development of secure automated management systems based on web technologies. *IOP Conf. Series: Materials Science and Engineering*, 2019, no. 537(5). doi:10.1088/1757-899X/537/5/052024
- Dikii D. I., Artemeva V. D. MQTT data protocol in remote access control management model for Internet networks. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2019, vol. 19, no. 1, pp. 109–117 (In Russian). doi:10.17586/2226-1494-2019-19-1-109-117
- Andy S., Rahardjo B., Hanindhito B. Attack scenarios and security analysis of MQTT communication protocol in IoT system. *4th Intern. Conf. on Electrical Engineering, Computer Science and Informatics*, 2017, pp. 1–6. doi:10.1109/EECSI.2017.8239179
- Tatarnikova T. M., Bogdanov P. Yu. Intrusion detection in internet of things networks based on machine learning methods. *Informatsionno-upravliaiushchie sistemy [Information and Control Systems]*, 2021, no. 6, pp. 42–52 (In Russian). doi:10.31799/1684-8853-2021-6-42-52
- Meidan Y., Bohadana M., Mathov Y., Mirsky Y., Breitenbacher D., Shabtai A., Elovici Y. N. BaIoT: Network-based detection of IoT botnet attacks using deep autoencoders. *IEEE Pervasive Computing, Special Issue – Securing the IoT*, 2018, no. 17(3), pp. 12–22.
- Duhamel P., Vetterli M. Fast Fourier transforms: a tutorial review and a state of the art. *Signal Processing*, 1990, no. 19, pp. 259–299.
- Kononovich E. V., Mironova I. V., Baturin V. A. Time-frequency analysis of solar activity series. *Investigated in Russia*, 2006, no. 182, pp. 1704–1715. Available at: (accessed 13 May 2022) (In Russian).
- Ziatdinov S. I. Reconstruction of signal by its samples on the base of Kotelnikov counts theorem. *Journal of Instrument Engineering*, 2010, no. 53(5), pp. 44–47 (In Russian).
- Lüke H. D. The origins of the sampling theorem. *IEEE Communications Magazine*, 1999, no. 37(4), pp. 106–108.



Разработка автоматизированного программного инструмента на основе методов машинного обучения для решения задач радиопланирования на участках метрополитена

А. А. Адёркина^а, аспирант, преподаватель, orcid.org/0000-0003-4331-4301, aaderkina@mail.ru

А. С. Синицын^а, студент, orcid.org/0000-0002-0793-950X

^аНациональный исследовательский Нижегородский государственный университет им. Н. И. Лобачевского, Гагарина пр., 23, Нижний Новгород, 603950, РФ

Введение: современный подход к радиопланированию обеспечивает пассажирам метрополитена бесперебойный доступ к сети Интернет благодаря использованию специальной модели распространения сигнала, которая вычисляет величины потерь мощности сигнала при его распространении между передатчиком и приемником на линиях метрополитена. Недостатком модели является высокая вычислительная сложность. **Цель:** разработать алгоритм предсказания величин потерь мощности сигнала с использованием методов машинного обучения, характеризующийся высокой точностью и низкой вычислительной сложностью. **Результаты:** анализ методов машинного обучения выявил, что максимально возможную точность при решении поставленной задачи обеспечивает метод случайного леса. Разработана структура данных для обучения выбранного метода и предсказания величин потерь мощности сигнала, содержащая параметры цифровой карты линий метрополитена. Сформирован ряд допущений при разработке итогового алгоритма: задача решается методом классификации, предсказываемые значения являются целыми числами. Разработан алгоритм предсказания величин потерь мощности сигнала, не использующий модель распространения напрямую, что снизило вычислительную сложность и уменьшило время выполнения задач радиопланирования при сохранении высокой точности предсказания. **Практическая значимость:** благодаря использованию методов машинного обучения в разработанных алгоритмах время выполнения задач радиопланирования сократилось с нескольких дней до нескольких часов с сохранением точности, что позволяет обрабатывать большее количество заказов по радиопланированию или же сокращать время работы инженеров на выполнение такого же количества заказов, что приводит к финансовой выгоде.

Ключевые слова — машинное обучение, потери мощности сигнала, алгоритмы предсказания, точность алгоритма, базовая станция, цифровая карта местности.

Для цитирования: Адёркина А. А., Синицын А. С. Разработка автоматизированного программного инструмента на основе методов машинного обучения для решения задач радиопланирования на участках метрополитена. *Информационно-управляющие системы*, 2022, № 5, с. 32–39. doi:10.31799/1684-8853-2022-5-32-39, EDN: IVCTMH

For citation: Aderkina A. A., Sinitsyn A. S. Development of an automated software tool based on machine learning methods for solving problems of radio planning in subway sections. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 5, pp. 32–39 (In Russian). doi:10.31799/1684-8853-2022-5-32-39, EDN: IVCTMH

Введение

Применение методов искусственного интеллекта в настоящее время широко распространено во многих сферах промышленности [1]. Целью данной работы является разработка алгоритма высокой точности для предсказания величины потери мощности радиосигнала при его распространении между передатчиком и приемником вдоль линий метрополитена. Разрабатываемый алгоритм направлен на снижение вычислительной сложности применяемой в настоящее время методики радиопланирования беспроводных сетей в метрополитене [2, 3].

Развитие технологий цифровой связи приводит к развертыванию все большего числа беспроводных сетей разных поколений и назначений. Неуклонно повышается роль беспроводной

связи в организации повседневной жизни людей, что привело к развертыванию зон покрытия сетей Wi-Fi (семейство стандартов IEEE 802.11) на общественных пространствах, в том числе внутри тоннелей и на открытых участках метрополитена. Сети такого типа состоят из мобильных станций, размещенных на крыше подвижного состава, и стационарно расположенных вдоль маршрута поезда базовых станций (БС). Качество развернутой сети в большой степени определяется правильным выбором точек расположения БС вдоль путей [4–8].

Основную трудность при выборе координат БС представляет расчет величины затухания сигнала (Path Loss, PL) между точками расположения базовых и мобильных станций в тоннелях метрополитена, так как тоннели имеют сложную геометрию с множеством поворотов,

спусков/подъемов и рассеивателей, которые препятствуют распространению сигнала. Для формирования подхода к определению наилучших точек расположения БС вдоль пути следования поездов разработана модель распространения сигнала в указанных условиях [2, 3]. На основе разработанной модели выработан комплексный подход к определению наилучшего расположения БС вдоль путей. Разработанный подход характеризуется высокой вычислительной сложностью. Для получения финальной расстановки БС требуется в некоторых случаях время до нескольких недель, что является существенным препятствием для развертывания беспроводных сетей на большом количестве участков, перепланирования сетей или сравнения нескольких вариантов конфигурации одной сети.

С целью устранить описанный недостаток существующего подхода предлагается разработка алгоритма на основе методов машинного обучения для предсказания уровня затухания мощности сигнала [9–12]. В работе применялись современные методы, позволившие не использовать напрямую модель распространения, что, в свою очередь, снизило время и стоимость оценки покрытия [9, 13–18].

В работах разных исследовательских групп изучались различные методы машинного обучения для предсказания величины PL [13, 18] или индикатора мощности принятого сигнала [14, 15, 17] для задач сетевого планирования в городской среде [13, 16–18] и внутри помещений [14]. Во всех изученных работах процедуры обучения и тестирования выполнялись с помощью набора данных о потерях на трассе, созданного на основе результатов моделирования сети, использующей цифровую модель местности [13, 16, 18], или на основе полевых измерений [14, 17].

В работе [13] исследовались и использовались следующие алгоритмы: опорный вектор регрессии (Support Vector Machine, SVM), случайный лес (Random Forest, RF) и К-ближайших соседей (K-Nearest Neighbor, KNN). Используемые результаты моделирования включали в себя оба условия распространения — при наличии прямой видимости и без. Результаты показывают, что все оцененные алгоритмы прогнозируют потери на трассе с высокой точностью, обеспечивая среднеквадратичные ошибки порядка 2–2,2 дБ при наличии прямой видимости и 3,4–4,1 дБ без соответственно. В процессе исследований алгоритмов KNN показал наилучшую производительность, таким образом, он является предпочтительным вариантом для прогнозирования потерь на трассе в городских условиях [13].

В работе [14] представлен инструмент оценки на основе машинного обучения, использующий не модель распространения [19], а измеренные

при эксплуатации сети величины индикатора мощности принятого сигнала, зависящие от расстояния, географии и параметров конфигурации БС [20]. Собранные данные применяются для обучения классификатора. После сравнения различных алгоритмов машинного обучения результаты показали, что метод классификации на основе SVM превосходит другие классификаторы, обеспечивая точность до 86,7 % [14].

Формирование данных для обучения и применения алгоритмов машинного обучения

Термин «машинное обучение» обозначает множество математических методов, решающих задачу поиска закономерностей между входными данными (признаками) и выходными данными (ответами) объекта [21]. Объектом в решаемой задаче является процесс затухания радиосигнала при распространении между точками передатчика и приемника вдоль линий метрополитена.

Целью разрабатываемого алгоритма является предсказание величины потерь мощности при распространении сигнала вдоль линий метрополитена по характеристикам геометрии участка распространения. Таким образом, в поставленной задаче признаками объекта являются геометрические параметры области распространения сигнала (цифровой карты зоны), а ответом — величина в децибелах потери мощности сигнала при распространении. В качестве среды распространения радиосигнала рассматриваются открытые участки и тоннели метрополитена.

Для характеристического описания области распространения задается ось линии метрополитена. Для упрощения расчетов ось задается двумя кривыми в горизонтальной и вертикальной плоскостях. Вдоль оси линии в каждой плоскости вводится криволинейная координатная ось. Также линия разбивается на набор сегментов, каждый из которых имеет следующие характеристики: ширина, высота, радиус поворота в горизонтальной плоскости, длина, направление поворота (лево/право или верх/низ), величина уклона в вертикальной плоскости.

Для составления цифровой карты зоны между фиксированными двумя точками участок между ними покрывается координатной сеткой следующим образом: координата одной из точек (передатчика или приемника) объявляется начальной, далее с выбранным шагом (например, 25 м) вдоль оси линии определяются следующие узлы сетки в сторону второй точки до момента превышения ее координат. Координата последней точки округляется до ближайшего узла построенной координатной сетки. В каждом узле

полученной координатной сетки определяются следующие параметры: радиус поворота и направление в горизонтальной плоскости, величина уклона и направление уклона в вертикальной плоскости, расстояние между точками передатчика (Tx) и приемника (Rx), тип рассматриваемого участка (тоннельный/наземный).

Минимально необходимым набором признаков для описания цифровой карты зоны распространения радиосигнала были выбраны следующие метрики:

- расстояния между точкам Tx и Rx вдоль оси линии;
- направление поворота пути в точках Tx, Rx в горизонтальной плоскости;
- радиус поворота пути в точках Tx, Rx в горизонтальной плоскости;
- направление уклона пути в точках Tx, Rx в вертикальной плоскости;
- величина уклона пути в точках Tx, Rx в вертикальной плоскости;
- идентификатор типа участка (открытый/тоннельный) в точках Tx, Rx;
- усредненные метрики между точками (направление поворота, радиус поворота, направление уклона, величина уклона).

В качестве дополнительных метрик рассматривались:

- медианные метрики между точками (направление поворота, радиус поворота, направление уклона, величина уклона);
- среднеквадратичные отклонения метрик между точками (направление поворота, радиус поворота, направление уклона, величина уклона) от средних значений;
- минимальное значение метрик между точками (направление поворота, радиус поворота, направление уклона, величина уклона);
- максимальное значение метрик между точками (направление поворота, радиус поворота, направление уклона, величина уклона).

Усредненные, медианные, среднеквадратичные, минимальные и максимальные значения находятя по массиву соответствующих метрик, полученных в каждом узле нанесенной на участок сетке.

Выбор алгоритма машинного обучения и набора признаков для предсказания величины потерь мощности сигнала

Выходной ответ любой задачи, решаемой методами машинного обучения, может быть или меткой одного из набора классов, тогда это задача классификации [22], или вещественным значением, тогда это задача регрессии [21]. При проведении данного исследования были рассмо-

трены наиболее популярные алгоритмы классификации и регрессии [22]. Исследуемые модели (табл. 1) могут быть использованы для решения задач обоих типов.

В данной работе исследованы и использованы методы машинного обучения с учителем, которые представляют собой набор математических функций, преобразующих входные в выходные данные на основе примеров пар ввода-вывода [20].

В первом подходе задача решалась как задача регрессии, т. е. выходными значениями были вещественные значения PL. Для оценки методов в этом случае использовалась метрика – коэффициент детерминации R^2 , равный отношению дисперсии предсказанных данных к дисперсии исходных данных [12]. Для разумных моделей коэффициент детерминации лежит в пределах от 0 до 1, причем случай $R^2 = 1$ соответствует идеальной модели, $R^2 = 0$ – модели на уровне оптимальной «константной», $R^2 < 0$ – модели хуже «константной» (такие алгоритмы не нужно рассматривать). Оптимальным константным алгоритмом называется такой алгоритм, который всегда возвращает среднее значение ответов для объектов обучающей выборки.

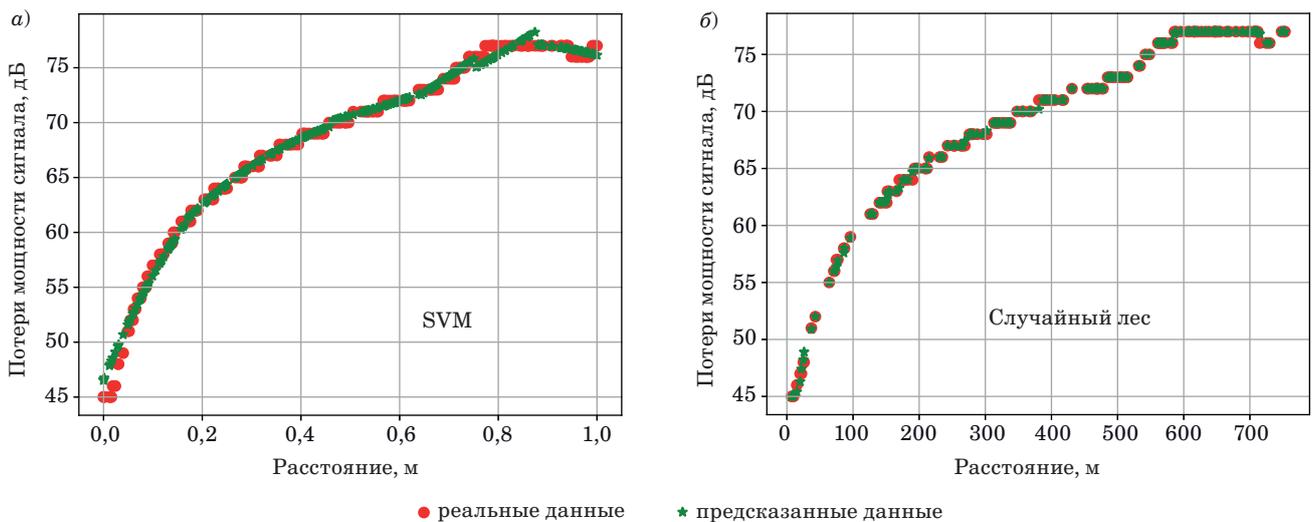
При решении исследуемой задачи как задачи классификации для оценки точности используется метрика точность, равная процентному отношению правильно предсказанных значений относительно всех значений.

Результаты применения обоих подходов представлены в табл. 1. Примеры зависимостей величины PL от расстояния между точками Tx и Rx, полученные с помощью исходной модели распространения сигнала [2, 3] и алгоритма на основе метода машинного обучения, приведены на рис. 1.

По значениям метрики R^2 , близким к единице, можно сделать вывод, что модели хорошо применимы для решения задач регрессии. На примере предсказания PL методом опорных векторов (рис. 1, а) видно, что предсказанные значения хорошо повторяют реальную зависимость. Но так как значения всех моделей, приведенных

■ **Таблица 1.** Сравнение моделей машинного обучения
 ■ **Table 1.** Comparison of machine learning models

Модель	Параметр R^2	Точность
SVM	0,99	0,71
Дерево решений	0,99	0,95
Случайный лес	0,99	0,98
К-ближайших соседей	0,99	0,94
Логистическая регрессия	1	0,8



■ **Рис. 1.** Демонстрация применения методов опорных векторов (а) и случайный лес (б) для предсказания величины PL
 ■ **Fig. 1.** Demonstration of the application of SVM (а) and Random Forest (б) methods for predicting the value of PL

в таблице, близки к единице, коэффициент детерминации невозможно использовать для сравнения методов между собой.

По значениям метрики точности при решении задачи классификации видно, что этот класс модели хорошо применим для решения исходной задачи. Эти выводы подтверждаются графически на рис. 1, б, где представлен набор реальных данных и предсказанных. В результате сравнения методов для решения задач проекта был выбран метод классификации случайный лес.

При исследовании возможности достижения максимальной точности в выбранном методе рассматривалось различное количество признаков в обучающих данных [12]. В первом подходе был взят набор из 16 минимально необходимых признаков, перечисленных в предыдущем разделе. Далее рассматривался расширенный набор из 32 метрик.

Использование расширенного набора метрик привело к незначительному повышению точности алгоритма на 0,015, но увеличило время выполнения в 1,8 раза. В результате было решено пользоваться набором из основных метрик.

Таким образом, при разработке алгоритма предсказания величины затухания мощности сигнала при его распространении между двумя фиксированными точками с помощью выбранного метода машинного обучения для достижения максимальной точности предсказания были сделаны следующие допущения:

- для разрабатываемого алгоритма выбран метод классификации случайный лес;
- для использования метода классификации выходные значения предсказанных величин выражались в целых значениях децибел.

Погрешность предсказания в таком подходе составит 1 дБ;

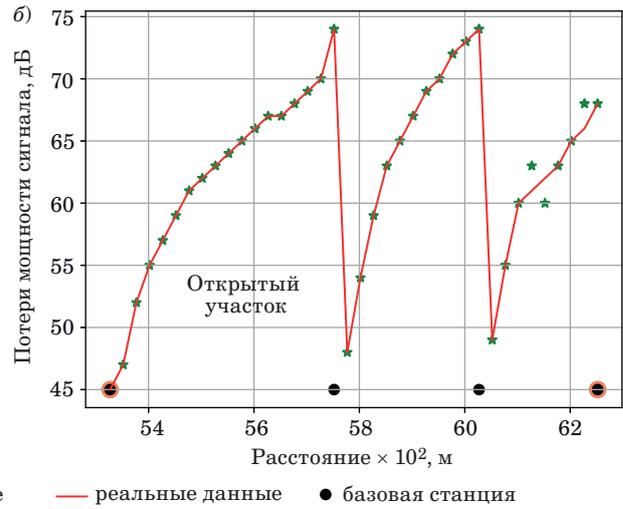
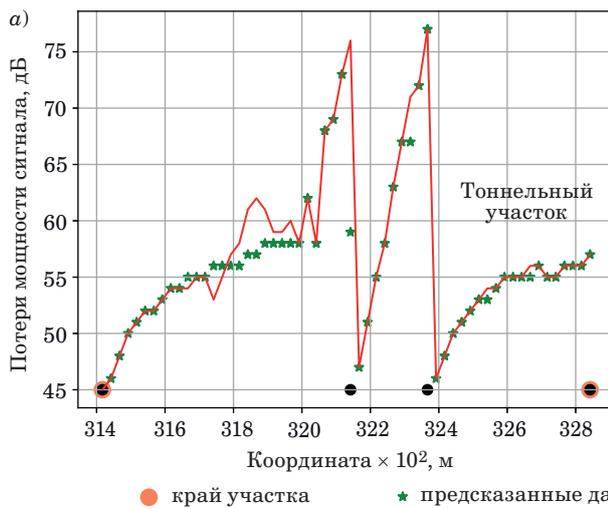
– для обучения разработанного алгоритма и предсказания величины PL использовался набор из 16 минимально необходимых признаков.

Примеры работы алгоритма предсказания PL и его точность

В ходе разработки алгоритма решена задача предсказания величины PL по значениям геометрических параметров области распространения. Примеры применения разработанного алгоритма представлены на рис. 2, а, б. Каждая точка на кривых – величина PL между соседней БС с меньшей координатой (черные точки на рисунках) и текущей точкой (например, значение PL для точки с координатой 31 600 м – это PL между БС с координатой 31 400 и текущей точкой 31 600). Дистанция между любыми двумя соседними точками каждой кривой составляет 25 м. По приведенным примерам видно соответствие предсказанных значений со значениями, предсказанными исходной моделью [2, 3].

Точность разработанного алгоритма приведена в табл. 2, где указаны средние значения ошибки разработанного алгоритма при предсказании PL и координат расположения БС по сравнению с существующим подходом на основе исходной модели распространения сигнала [2].

Значения, приведенные в табл. 2, показывают, что разработанный алгоритм характеризуется высокой точностью, а главным его преимуществом является существенное снижение времени выполнения. Исходный алгоритм радиоплани-



■ **Рис. 2.** Применение разработанного алгоритма на тоннельном (а) и открытом (б) участке
 ■ **Fig. 2.** Application of the developed algorithm in the tunnel (а) and open (б) section

■ **Таблица 2.** Точность разработанного алгоритма
 ■ **Table 2.** Accuracy of the developed algorithm

Параметр	Значение сценария	
	тоннельного	наземного
Ошибка предсказания PL, дБ	1,9	0,3
Ошибка определения координат БС, м	107	75
Ошибка плотности БС, БС/км	0,5	0,3

рования [2] требует многократного вызова исходной модели распространения сигнала. Время выполнения радиопланирования, заключающегося в расстановке БС на одном тоннельном участке длиной 3 км линии метрополитена, составляет около часа. С помощью разработанного алгоритма предсказания PL, использующего метод случайного леса, удалось сократить время выполнения до одной минуты.

Заключение

Проведены анализ методов машинного обучения и сравнение методов регрессии и классификации; определен метод, обеспечивающий максимально возможную точность при решении поставленной задачи предсказания величины PL; разработана структура данных для обучения выбранного метода. На основании информации о цифровой карте исследуемого участка сформирован ряд допущений при разработке итогового

алгоритма предсказания величины потери мощности сигнала.

Разработанный алгоритм предсказания PL характеризуется высокой точностью и малым временем исполнения, что позволило существенно сократить время выполнения радиопланирования.

Финансовая поддержка

Работа выполнена за счет гранта Фонда содействия развитию малых форм предприятий в научно-технической сфере (№ 16184ГУ/2020) в рамках проекта «Разработка методов машинного обучения при определении оптимального количества и расположения базовых станций для обеспечения непрерывного покрытия тоннелей общественного транспорта радиосетью».

Литература

1. Тюльпа Д. Применение искусственного интеллекта в промышленности. Предпосылки и возможности. *САПР и графика*, 2020, № 1, с. 26–30. <https://sapr.ru/article/25989> (дата обращения: 15.05.2021).
2. Адёркина А. А., Шашанов М. А., Шумилов В. Ю., Трушанин А. Ю. Радиопланирование систем беспроводной связи в тоннелях метрополитена. *Радиотехнические и телекоммуникационные системы*, 2020, № 1, с. 41–53.
3. Адёркина А. А., Шашанов М. А., Шумилов В. Ю., Трушанин А. Ю. Измерение и анализ ослабления сигнала диапазона 5 ГГц при распространении вдоль участков метрополитена различного типа.

Радиотехнические и телекоммуникационные системы, 2021, № 1, с. 33–45.

4. **Domínguez-Bolaño T., Rodríguez-Piñeiro J., García-Nay J. A., Yin X., Castedo L.** Measurement-based characterization of train-to-infrastructure 2.6 GHz propagation channel in a modern subway station. *IEEE Access*, 2018, vol. 6, pp. 52814–52830. doi:10.1109/ACCESS.2018.2870564
5. **Briso-Rodríguez C., Fratilesco P., Xu Y.** Path loss modeling for train-to-train communications in subway tunnels at 900/2400 MHz. *IEEE Antennas and Wireless Propagation Letters*, 2019, vol. 18, no. 6, pp. 1164–1168. doi:10.1109/LAWP.2019.2911406
6. **Guan K., Zhong Z., Ai B., Briso-Rodríguez C.** Statistical modeling for propagation in tunnels based on distributed antenna systems. *IEEE Antennas and Propagation Society International Symposium (AP-SURSI)*, 2013, pp. 1920–1921. doi:10.1109/APS.2013.6711618
7. **Hou X., Shu H.** Research on radio network planning and optimization in TETRA digital trunking system. *IEEE Second Intern. Conf. on Instrumentation, Measurement, Computer, Communication and Control*, 2012, pp. 821–824. doi:10.1109/IMCCC.2012.198
8. **Roullier-Callaghan A.** A radio coverage and planning tool. *6th IEEE High Frequency Postgraduate Colloquium (Cat. No. 01TH8574)*, 2001, pp. 35–40. doi:10.1109/HFPSC.2001.962156
9. **Binzer T., Landstorfer F. M.** Radio network planning with neural networks. *Vehicular Technology Conf. Fall 2000. IEEE VTS Fall VTC2000. 52nd Vehicular Technology Conf. (Cat. No. 00CH37152)*, 2000, vol. 2, pp. 811–817. <https://ieeexplore.ieee.org/document/887116> (дата обращения: 15.08.2021).
10. **Van Nguyen T., Jeong Y., Shin H., Win M. Z.** Machine learning for wideband localization. *IEEE Journal on Selected Areas in Communications*, 2015, vol. 33, no. 7, pp. 1357–1380. doi:10.1109/JSAC.2015.2430191
11. **Alvizu R., Troia S., Maier G., Pattavina A.** Matheuristic with machine-learning-based prediction for software-defined mobile metro-core networks. *Journal of Optical Communications and Networking*, 2017, vol. 9, no. 9, pp. D19–D30.
12. **VanderPlas J.** *Python data science handbook: Essential tools for working with data.* O'Reilly Media, Inc., 2017. 546 p.
13. **Moraitis N., Tsipi L., Vouyioukas D.** Machine learning-based methods for path loss prediction in urban environment for LTE networks. *16th Intern. Conf. on Wireless and Mobile Computing, Networking and Communications (WiMob), Thessaloniki (Greece)*, 2020. <https://ieeexplore.ieee.org/document/9253369> (дата обращения: 15.08.2021).
14. **Dai L., Zhang H., Zhuang Y.** Propagation-model-free coverage evaluation via machine learning for future 5G networks. *IEEE 29th Annual Intern. Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2018. <https://ieeexplore.ieee.org/document/8580992> (дата обращения: 15.08.2021).
15. **Moysen J., Giupponi L., Mangues-Bafalluy J.** A machine learning enabled network planning tool. *IEEE 27th Annual Intern. Symp. on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2016, pp. 1–7.
16. **Abdallah A. A., Saab S. S., Kassas Z. M.** A machine learning approach for localization in cellular environments. *IEEE/ION Position, Location and Navigation Symp. (PLANS)*, 2018, pp. 1223–1227.
17. **Politanskyi R., Klymash M.** Application of artificial intelligence in cognitive radio for planning distribution of frequency channels. *3rd Intern. Conf. on Advanced Information and Communications Technologies (AICT)*, 2019, pp. 390–394.
18. **Vallejo-Cabrejas G., Batolome-Pascual P.** Application of radio channel modelling to a planning tool in a mobile radio indoor communication system. *IEEE 43rd Vehicular Technology Conf.*, 1993, pp. 219–222.
19. **Dai L., Zhang H.** Propagation-model-free base station deployment for mobile networks: Integrating machine learning and heuristic methods. *IEEE Access*, 2020, vol. 8, pp. 83375–83386. doi:10.1109/ACCESS.2020.2990631
20. **Raj N.** Indoor RSSI prediction using machine learning for wireless networks. *Intern. Conf. on COMMunication Systems & NETWORKS (COMSNETS)*, 2021, pp. 372–374. <https://ieeexplore.ieee.org/document/9352852> (дата обращения: 15.08.2021).
21. **Müller A. C., Guido S.** *Introduction to machine learning with Python: A guide for data scientists.* O'Reilly Media, Inc., 2016. 398 p.
22. **Le J.** *The top 10 machine learning algorithms every beginner should know.* https://builtin.com/data-science/tour-top-10-algorithms-machine-learning-newbies?utm_campaign=News&utm_medium=Community&utm_source=DataCamp.com (дата обращения: 05.02.2021).

UDC 004.021, 004.896, 004.852
 doi:10.31799/1684-8853-2022-5-32-39
 EDN: IVCTMH

Development of an automated software tool based on machine learning methods for solving problems of radio planning in subway sections

A. A. Aderkina^a, Post-Graduate Student, Lecturer, orcid.org/0000-0003-4331-4301, aaderkina@mail.ru
 A. S. Sinitsyn^a, Student, orcid.org/0000-0002-0793-950X

^aNational Research Lobachevsky State University of Nizhny Novgorod, 23, Gagarin Ave., 603950, Nizhny Novgorod, Russian Federation

Introduction: The modern approach to radio planning provides subway passengers with uninterrupted access to the Internet. This is achieved through the use of a special signal propagation model which calculates signal power loss during its propagation between a transmitter and a receiver on subway lines. The disadvantage of the model is the high computational complexity. **Purpose:** Using machine learning methods to develop an algorithm for predicting the signal power loss, the algorithm being characterized by high accuracy and low computational complexity. **Results:** The analysis of machine learning methods revealed that the maximum possible accuracy in solving the problem is provided by the random forest method. A data structure containing the parameters of a digital map of subway lines was developed to train the selected method and predict a signal power loss. While developing the final algorithm a number of assumptions were made, such as: the problem is solved as a classification problem, the predicted values are integers. A signal power loss prediction algorithm that does not directly use the propagation model was developed, which reduced the computational complexity and the execution time for solving radio planning problems, with high prediction accuracy maintained. **Practical relevance:** Due to the use of machine learning methods in developed algorithms the time for performing radio planning was reduced from several days to several hours, with accuracy preserved. This allows to process more radio planning orders or to reduce the working time for engineers to complete the same number of orders, which is a financial benefit.

Keywords – machine learning, signal power loss, prediction algorithms, accuracy of an algorithm, base station, digital map.

For citation: Aderkina A. A., Sinitsyn A. S. Development of an automated software tool based on machine learning methods for solving problems of radio planning in subway sections. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 5, pp. 32–39 (In Russian). doi:10.31799/1684-8853-2022-5-32-39, EDN: IVCTMH

Financial support

The work was supported by a grant from the Foundation for the Promotion of the Development of Small Forms of Enterprises in the Scientific and Technical Sphere (No. 16184GU/2020) as part of the project “Development of machine learning methods for determining the optimal number and location of base stations to ensure continuous coverage of public transport tunnels with a radio network”.

Reference

1. Tyulpa D. The use of artificial intelligence in industry. Prerequisites and opportunities. *SAPR i grafika*, 2020, no. 1, pp. 26–30. Available at: <https://sapr.ru/article/25989> (accessed 15 May 2021) (In Russian).
2. Aderkina A. A., Shashanov M. A., Shumilov V. Yu., Trushanin A. Yu. Radio planning of wireless communication systems in metro tunnels. *Radio Engineering and Telecommunication Systems*, 2020, no. 1, pp. 41–53 (In Russian).
3. Aderkina A. A., Shashanov M. A., Shumilov V. Yu., Trushanin A. Yu. Measurement and analysis of signal attenuation in the 5 GHz band during propagation along various types of subway sections. *Radio Engineering and Telecommunication Systems*, 2021, no. 1, pp. 33–45 (In Russian).
4. Domínguez-Bolaño T., Rodríguez-Piñero J., García-Nay J. A., Yin X., Castedo L. Measurement-based characterization of train-to-infrastructure 2.6 GHz propagation channel in a modern subway station. *IEEE Access*, 2018, vol. 6, pp. 52814–52830. doi:10.1109/ACCESS.2018.2870564
5. Briso-Rodríguez C., Fratilesco P., Xu Y. Path loss modeling for train-to-train communications in subway tunnels at 900/2400 MHz. *IEEE Antennas and Wireless Propagation Letters*, 2019, vol. 18, no. 6, pp. 1164–1168. doi:10.1109/LAWP.2019.2911406
6. Guan K., Zhong Z., Ai B., Briso-Rodríguez C. Statistic modeling for propagation in tunnels based on distributed antenna systems. *IEEE Antennas and Propagation Society Intern. Symp. (APSURSI)*, 2013, pp. 1920–1921. doi:10.1109/APS.2013.6711618
7. Hou X., Shu H. Research on radio network planning and optimization in TETRA digital trunking system. *IEEE Second Intern. Conf. on Instrumentation, Measurement, Computer, Communication and Control*, 2012, pp. 821–824. doi:10.1109/IMCCC.2012.198
8. Roullier-Callaghan A. A radio coverage and planning tool. *6th IEEE High Frequency Postgraduate Colloquium (Cat. No. 01TH8574)*, 2001, pp. 35–40. doi:10.1109/HFPSC.2001.962156
9. Binzer T., Landstorfer F. M. Radio network planning with neural networks. *Vehicular Technology Conf. Fall 2000. IEEE VTS Fall VTC2000. 52nd Vehicular Technology Conf. (Cat. No. 00CH37152)*, 2000, vol. 2, pp. 811–817. Available at: <https://ieeexplore.ieee.org/document/887116> (accessed 15 August 2021).
10. Van Nguyen T., Jeong Y., Shin H., Win M. Z. Machine learning for wideband localization. *IEEE Journal on Selected Areas in Communications*, 2015, vol. 33, no. 7, pp. 1357–1380. doi:10.1109/JSAC.2015.2430191
11. Alvizu R., Troia S., Maier G., Pattavina A. Matheuristic with machine-learning-based prediction for software-defined mobile metro-core networks. *Journal of Optical Communications and Networking*, 2017, vol. 9, no. 9, pp. D19–D30.
12. VanderPlas J. *Python data science handbook: Essential tools for working with data*. O'Reilly Media, Inc., 2017. 546 p.
13. Moraitis N., Tsipi L., Vouyioukas D. Machine learning-based methods for path loss prediction in urban environment for LTE networks. *16th Intern. Conf. on Wireless and Mobile Computing, Networking and Communications (WiMob), Thessaloniki (Greece)*, 2020. Available at: <https://ieeexplore.ieee.org/document/9253369> (accessed 15 August 2021).
14. Dai L., Zhang H., Zhuang Y. Propagation-model-free coverage evaluation via machine learning for future 5G networks. *IEEE 29th Annual Intern. Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2018. Available at: <https://ieeexplore.ieee.org/document/8580992> (accessed 15 August 2021).
15. Moysen J., Giupponi L., Mangles-Bafalluy J. A machine learning enabled network planning tool. *IEEE 27th Annual Intern. Symp. on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2016, pp. 1–7.
16. Abdallah A. A., Saab S. S., Kassas Z. M. A machine learning approach for localization in cellular environments. *IEEE/ION Position, Location and Navigation Symp. (PLANS)*, 2018, pp. 1223–1227.
17. Politansky R., Klymash M. Application of artificial intelligence in cognitive radio for planning distribution of frequency channels. *3rd Intern. Conf. on Advanced Information and Communications Technologies (AICT)*, 2019, pp. 390–394.

18. Vallejo-Cabrejas G., Batolome-Pascual P. Application of radio channel modelling to a planning tool in a mobile radio indoor communication system. *IEEE 43rd Vehicular Technology Conf.*, 1993, pp. 219–222.
 19. Dai L., Zhang H. Propagation-model-free base station deployment for mobile networks: Integrating machine learning and heuristic methods. *IEEE Access*, 2020, vol. 8, pp. 83375–83386. doi:10.1109/ACCESS.2020.2990631
 20. Raj N. Indoor RSSI prediction using machine learning for wireless networks. *Intern. Conf. on COMMunication Systems & NETworkS (COMSNETS)*, 2021, pp. 372–374. Available at: <https://ieeexplore.ieee.org/document/9352852> (accessed 15 August 2021).
 21. Müller A. C., Guido S. *Introduction to machine learning with Python: A guide for data scientists*. O'Reilly Media, Inc., 2016. 398 p.
 22. Le J. *The top 10 machine learning algorithms every beginner should know*. Available at: https://builtin.com/data-science/tour-top-10-algorithms-machine-learning-newbies?utm_campaign=News&utm_medium=Community&utm_source=DataCamp.com (accessed 5 February 2021).
-

ПАМЯТКА ДЛЯ АВТОРОВ

Поступающие в редакцию статьи проходят обязательное рецензирование.

При наличии положительной рецензии статья рассматривается редакционной коллегией. Принятая в печать статья направляется автору для согласования редакторских правок. После согласования автор представляет в редакцию окончательный вариант текста статьи.

Процедуры согласования текста статьи могут осуществляться как непосредственно в редакции, так и по e-mail (ius.spb@gmail.com).

При отклонении статьи редакция представляет автору мотивированное заключение и рецензию, при необходимости доработать статью — рецензию.

Редакция журнала напоминает, что ответственность за достоверность и точность рекламных материалов несут рекламодатели.



Построение архитектуры туманных вычислений с использованием технологии блокчейн

А. В. Пименов^а, студент, orcid.org/0000-0002-9136-3514

И. Р. Федоров^а, аспирант, orcid.org/0000-0003-2422-4714

С. В. Беззатеев^б, доктор техн. наук, профессор, orcid.org/0000-0002-0924-6221, bezzateev_sergey@mail.ru

^аУниверситет ИТМО, Кронверкский пр., 49, Санкт-Петербург, 197101, РФ

^бСанкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ

Введение: по мере роста количества и многообразия устройств, подключенных к сети Интернет, растут требования как к производительности сети, так и к обеспечению безопасности хранения и передачи данных. Проблемы производительности решают, как правило, за счет облачных, туманных и граничных вычислений, а проблема безопасности хранения и передачи данных остается актуальной. Одним из эффективных путей ее решения является использование технологии блокчейн. **Цель:** проектирование архитектуры сети туманных вычислений на основе технологии блокчейн. **Результаты:** проведенные исследования в области туманных вычислений позволили определить требования к их архитектуре: автономность, масштабируемость, гибкость, иерархичность, безопасность, надежность, доступность, удобство обслуживания. Выделенные критерии построения архитектуры обусловили выбор в пользу частного блокчейна из-за его более высокой производительности по сравнению с открытым блокчейном. Проведен сравнительный анализ алгоритмов консенсуса, которые чаще других используются в частных блокчейнах, и выбран наиболее подходящий. В соответствии с установленными требованиями и результатами анализа спроектирована модель архитектуры туманных вычислений на основе частного блокчейна. Архитектура включает четыре элемента: конечные устройства, туманные узлы, узлы оркестрации и облачную инфраструктуру. В блокчейн входят туманные узлы и узлы оркестрации, за счет чего обеспечивается конфиденциальность, доступность и целостность данных в туманной сети. **Практическая значимость:** результаты исследования могут быть использованы при проектировании сетей туманных вычислений как по отдельности, так и в составе мобильных сетей 5G.

Ключевые слова — туманные вычисления, архитектура туманных вычислений, блокчейн, информационная безопасность, оркестрация, интернет вещей.

Для цитирования: Пименов А. В., Федоров И. Р., Беззатеев С. В. Построение архитектуры туманных вычислений с использованием технологии блокчейн. *Информационно-управляющие системы*, 2022, № 5, с. 40–48. doi:10.31799/1684-8853-2022-5-40-48, EDN: KJPXLT

For citation: Pimenov A. V., Fedorov I. R., Bezzateev S. V. Designing fog computing architecture with the use of blockchain technology. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 5, pp. 40–48 (In Russian). doi:10.31799/1684-8853-2022-5-40-48, EDN: KJPXLT

Введение

В современном мире связь приобретает ключевое значение, и ее технологии быстро развиваются. На данный момент к сетям связи подключены миллиарды различных устройств, а к 2030 году даются прогнозы на десятки миллиардов подключенных устройств [1]. Существуют такие технологии, как интернет вещей (Internet of Things, IoT), которые предъявляют высокие требования к связи между устройствами. Общемировая цифровизация требует огромного количества используемых устройств, имеющих выход как минимум в локальную сеть, а зачастую и в облако [2].

С развитием технологии IoT во всех сферах возникают следующие проблемы: недостаток на устройствах вычислительных ресурсов и информационная безопасность (ИБ) данных, содержащихся в сети. Если проблему недостатка

мощностей сегодня решают в основном за счет облачных, туманных и граничных вычислений, то проблема безопасности IoT-устройств стоит очень остро. Согласно данным ENISA (Агентство Европейского союза по сетям и информационной безопасности), инциденты с устройствами интернета вещей входят в тройку угроз с наибольшим финансовым ущербом для компаний.

На данный момент уже есть различные предложения по архитектуре туманных вычислений, однако они сталкиваются с определенными вопросами в области обеспечения ИБ данных и оптимизации потребления ресурсов у устройств, предоставляющих свои услуги для туманных вычислений. Эталонная архитектура OpenFog RA [3] предлагает решение проблем, однако не все варианты этих решений рассмотрены.

В 2018 году Национальный институт стандартов и технологий США сформулировал официальное определение термина туманные вычисления:

«Туманные вычисления – это многоуровневая модель, обеспечивающая повсеместный доступ к общей совокупности масштабируемых вычислительных ресурсов. Туманные узлы являются контекстно-зависимыми и поддерживают единую систему управления данными и организации связи. Туманные вычисления минимизируют время сетевого отклика поддерживаемых приложений, а также обеспечивают конечные устройства локальными вычислительными ресурсами и, при необходимости, сетевым подключением к централизованным сервисам».

В туманных вычислениях можно выделить следующие проблемы и угрозы ИБ [4–6].

1. Проблема аутентификации устройств. Чтобы получить доступ к службам сети тумана, устройство должно сначала стать частью сети, пройдя аутентификацию. Это представляет собой серьезную проблему, поскольку устройства ограничены различными параметрами.

2. Проблема безопасности беспроводной передачи данных. Обеспечение безопасности туманной сети затруднено из-за уязвимостей стандартных протоколов беспроводной передачи данных.

3. Проблема, связанная с доверием устройств внутри в сети. Устройства туманных вычислений часто разворачиваются без строгого контроля и защиты, поэтому они подвержены всем типам угроз безопасности.

4. Проблема конфиденциальности конечных пользователей. Узлы тумана находятся в непосредственной близости от конечных пользователей и могут собирать конфиденциальные данные.

5. Вредоносные атаки. Среда туманных вычислений может подвергаться многочисленным вредоносным атакам, и, таким образом, без удобных мер безопасности возможности сети могут быть серьезно подорваны.

Перечисленные проблемы не позволяют гарантировать конфиденциальность и целостность информации в туманных сетях. Это может привести к раскрытию конфиденциальных данных пользователей, сбою операций в сети, репутационным и материальным потерям корпораций. В данной работе мы рассматриваем возможность решения выявленных проблем с помощью технологии блокчейн.

Преимущества интеграции технологии блокчейн и туманных вычислений

Блокчейн – это технология хранения данных в цепочке последовательно связанных блоков. Каждый блок содержит уникальный код, называемый хешем. Блок также содержит хеш предыдущего блока в цепочке. После того как запись

добавлена в цепочку, ее невозможно изменить. Несмотря на то, что блокчейн подвержен специфическим атакам, его широко используют при реализации различных проектов для решения актуальных проблем ИБ.

1. Защита от DDoS. Предполагается создание одноранговой сети доставки контента на основе неиспользуемой полосы пропускания участников сети. В проекте использовалась инфраструктура блокчейна Ethereum и предполагалась экономическая модель мотивации подключенных участников на основе собственного цифрового актива [7].

2. Безопасность конечных устройств интернета вещей. Регистрация устройства в доверенном распределенном реестре теоретически позволяет устранить различные уязвимости интернета вещей (обеспечение целостности прошивок устройств, проблемы с подключением и аутентификацией). Одним из базовых механизмов защиты устройств с помощью технологии блокчейн является хранение контрольной суммы метаданных ПО прошивки (версия, время обновления и т. п.) в блокчейне. Любое обновление сверяется с доверенным журналом-протоколом изменений, тем самым обеспечивается гарантия целостности устройства. На сегодня защита распределенных сетей является одним из наиболее перспективных прикладных применений технологии блокчейн в сфере ИБ. Стоит также отметить реализации распределенного реестра, нацеленные на защиту IoT (проект IOTA) [8].

3. Децентрализованная идентификация и аутентификация. Концепция децентрализованной идентификации с помощью блокчейна предполагает, что пользователи могут самостоятельно хранить свои персональные данные. Таким образом обеспечивается полный контроль над доступом к личной информации [9].

В целом можно выделить следующие основные преимущества от интеграции блокчейна и туманных вычислений.

1. Блокчейн позволяет обеспечивать целостность, доступность и конфиденциальность, т. е. ИБ данных в туманной сети.

2. Блокчейн решает проблему обеспечения аутентификации устройств в туманной сети.

3. Использование блокчейна повышает защиту сети от DoS, вредоносных и других атак, а также от помех и искажений самой сети. Конечно, блокчейн имеет свои уязвимости, однако использование частного блокчейна с нужным алгоритмом консенсуса позволяет нивелировать большинство из них.

4. Блокчейн позволяет обеспечивать постоянный мониторинг используемых и доступных ресурсов в каждом кластере сети или узле, т. е. повышается осведомленность внутри сети.

5. Участники туманных вычислений могут получать вознаграждение за счет предоставления своих ресурсов туманной сети, что создает мотивацию для участия в туманных вычислениях.

Проект модели туманных вычислений с использованием технологии блокчейн

Больше всего применений туманные вычисления находят в «умных» системах, которые требуют обработки информации в режиме реального времени: умные автомобили, дроны-доставщики, умные дома [10]. Все это можно объединить в рамках умного города, поэтому при проектировании модели будем отталкиваться от этого объекта.

Также надо иметь в виду уже выполненные исследования в области проектирования архитектуры туманных вычислений: OpenFog RA, NIST [11] и др. В ходе исследований мы использовали некоторые предложенные решения в построении архитектуры и придерживались принципов, определенных в OpenFog:

- автономность;
- масштабируемость;
- открытость;
- безопасность;
- RAS (надежность, доступность, удобство обслуживания);
- гибкость;
- иерархичность;
- программируемость.

Выделим основные употребляемые нами элементы.

1. Конечное устройство – основной клиент туманных служб. Может представлять собой умную машину, роутер, компьютер. Надо понимать, что в парадигме туманных вычислений датчики в туманной машине не обращаются напрямую в туман, а участвуют в граничных вычислениях, а сами граничные вычисления по необходимости могут обращаться в туман или облако. Одно конечное устройство могут обслуживать множество туманных узлов.

2. Туманный узел – основной элемент туманной сети. Он представляет собой физическое или виртуальное устройство, которое предоставляет вычислительные услуги туманной сети. Модели их развертывания такие же, как и у облачных вычислений.

3. Узел оркестрации – управляющий элемент в туманной сети. Он организует работу туманных узлов, ведет учет и мониторинг ресурсов, управляет событиями и имеет прочие организующие функции, необходимые для нормального функционирования туманной сети. Каждый узел управляет одним гибким сегментом туманной сети.

4. Облачная инфраструктура – туманные вычисления разрабатывались как дополнение к облаку, поэтому архитектура туманных вычислений должна учитывать связь с облачной инфраструктурой для различных «тяжелых» вычислений или хранения большого количества информации.

Модель открытого блокчейна не подходит по многим причинам. Во-первых, она не обеспечивает соблюдение всех аспектов ИБ данных, а именно конфиденциальности. В туманных вычислениях может содержаться множество конфиденциальной информации, к тому же она обрабатывается близко к конечным устройствам, что может позволить установить личность владельца. Во-вторых, открытый блокчейн требует много ресурсов на подтверждение транзакций, что критично для туманных вычислений, так как они должны поддерживать системы реального времени. В проведенном исследовании [12] производительности открытого блокчейна Ethereum и частного блокчейна Hyperledger Fabric доказано, что производительность частного блокчейна выше в несколько десятков раз и может быть выше даже в сотни раз. Еще у открытых блокчейнов существует проблема масштабируемости, что также критично для туманных вычислений. В-третьих, мы не можем полностью доверять туманным узлам, так как не можем контролировать их развертывание. Даже авторизованное устройство может использоваться злоумышленником для атаки на туманную сеть.

Это приводит к тому, что требуется рассматривать архитектуру частных или гибридных блокчейнов. Использование частного блокчейна по сравнению с открытым блокчейном несет следующие преимущества.

1. Конфиденциальность. Используя частный блокчейн, можно быть уверенным, что передаваемая и хранимая информация в туманных сетях будет иметь аспект конфиденциальности ИБ данных.

2. Низкое потребление ресурсов. Алгоритмы консенсуса у частного блокчейна требуют меньшего количества участия узлов, что приводит к снижению потребления пропускной способности у узлов и уменьшению накладных расходов.

3. Обеспечение доверия внутри сети. Главные узлы (узлы оркестрации) будут разворачиваться под контролем, что повышает уровень доверия внутри сети.

Использование гибридного блокчейна в рамках данной статьи не будем затрагивать, так как, несмотря на его преимущества с точки зрения распределения прав, он требует подробного изучения и анализа в качестве платформы для взаимодействия устройств в туманных вычислениях. Прежде чем приступить к его изучению, следует

рассмотреть варианты построения архитектуры с помощью частного блокчейна, и, уже зная ее недостатки, можно устранять их с помощью гибридного блокчейна.

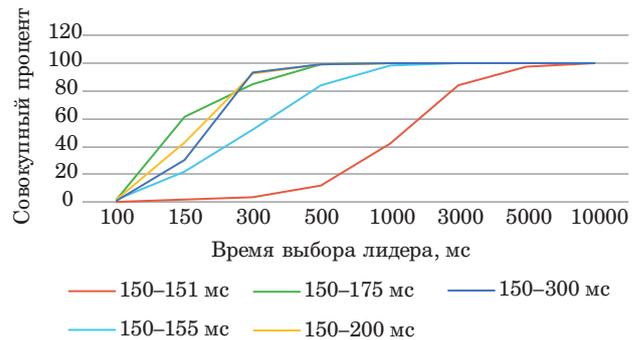
Второй немаловажной деталью является алгоритм консенсуса. На данный момент не существует специализированного алгоритма консенсуса для туманных вычислений. В настоящей работе рассмотрены следующие алгоритмы консенсуса: PBFT, PoAh, PoA и Raft. Выбор этих алгоритмов обусловлен их распространением в исследованиях, связанных с алгоритмами консенсусов частных блокчейнов и архитектур для IoT, граничных и туманных вычислений [13–17]. В рамках данного исследования не будем подробно останавливаться на принципах работы алгоритмов, а рассмотрим их специфику, производительность, преимущества и недостатки. Проведя сравнительный анализ, выберем алгоритм или предложим альтернативу на основе предложенных алгоритмов.

Алгоритм PBFT является одним из основных алгоритмов для частных блокчейнов. Популярная блокчейн-платформа Hyperledger Fabric использует этот алгоритм консенсуса, что дает высокую эффективность, устойчивость к сбоям и обеспечивает работу до участия трети неисправных узлов. Главным недостатком PBFT является низкая масштабируемость.

Алгоритм консенсуса PoA — довольно популярное решение среди различных предприятий и компаний. В нем каждый валидатор является абсолютно доверенным узлом за счет его разворачивания доверенным участником. Главным преимуществом алгоритма PoA является более высокая эффективность по сравнению с PoW, простота и масштабируемость [18, 19]. Из недостатков можно выделить слабую производительность по сравнению с более производительными алгоритмами и отсутствие мотивации для рядовых участников в блокчейне. Невозможность задать мотивацию является достаточно серьезным недостатком, который будет рассмотрен ниже.

Одним из алгоритмов, который тоже получил широкое распространение, является Raft. Его особенностью является то, что в блокчейне присутствует выборный лидер и его подписчики, которые синхронизируют свои узлы с ним. Это обеспечивает простоту, надежность, высокую эффективность (решения принимает один узел-лидер) и масштабируемость, однако требует дополнительных задержек и ресурсов на проведение выборов (рис. 1).

Кроме широко используемых алгоритмов консенсуса, рассмотрим алгоритм PoAh, который был предложен для применения в крупных сетях IoT. Этот алгоритм предназначен для решения проблемы высокого потребления ресурсов иными



■ **Рис. 1.** Анализ достижения выбора лидера при различных значениях тайм-аута

■ **Fig. 1.** Analysis of achieving leader election at different timeout values

алгоритмами консенсуса в маломощных устройствах IoT. Он содержит механизм аутентификации с цифровой подписью для доступа и проверки блока. В нем также вводятся уровни доверия, которые имеют базовое значение и меняются со временем в зависимости от качества работы узла. Соответственно, доверенные узлы, которые могут доказать свою идентичность, участвуют в консенсусе. Такой подход существенно снижает расходы за счет того, что современные асимметричные методы шифрования очень быстры. Несмотря на то, что применение PoAh больше подходит для граничных вычислений, где участвуют преимущественно маломощные устройства, идеи этого алгоритма консенсуса тоже можно использовать в туманных вычислениях.

Результаты исследований внесены в таблицу.

Как видно из таблицы, в исследованиях проводился анализ различных частей того или иного алгоритма консенсуса (в Raft исследуется скорость выбора лидера, а, например, в PoAh анализируется скорость аутентификации устройства и ответа от блокчейна). Кроме того, видны различия по характеристикам и архитектуре среды тестирования, что несколько усложняет сравнение различных алгоритмов, поэтому при выборе алгоритма консенсуса будем отталкиваться от потребностей нашей архитектуры.

Выделим основные требования, которые должны обеспечивать туманные вычисления в рамках умного города:

- 1) поддержка систем реального времени;
- 2) оказание услуг по обработке и хранению данных;
- 3) обеспечение высокого уровня ИБ данных в туманной сети;
- 4) мотивация рядовых участников для предоставления ресурсов туманным вычислениям;
- 5) реализация принципов гибкости, масштабируемости и иерархичности сети;

- Результат анализа производительности алгоритмов консенсуса
- The result of the consensus algorithms' performance analysis

Алгоритм консенсуса	Производительность консенсуса	Характеристики макета (виртуальной машины)	Комментарии
PBFT	Максимальное TPS: 45 Задержка для 50 транзакций: 1,5 с	Intel i7-7700k, ОЗУ 32 ГБ, жесткий диск 512 SSD. Построен в Hyperledger Fabric	Анализ производительности в исследованиях дал примерно одинаковый результат даже при различных характеристиках испытательного стенда, так как блокчейн не потребляет все ресурсы в ОС
	Максимальное TPS: 52 Задержка для 50 транзакций: 1,29 с	4 ядра Intel Cascade Lake, 8 Гб оперативной памяти, ОС Ubuntu 20 LTS, 20 Гб HDD. Построен в Hyperledger Fabric	
PoAh	Создание, заполнение транзакциями и интегрирование блока размером 35 байт за 3,34 с	Шесть одноплатных компьютеров Raspberry Pi	Для анализа производительности применялись маломощные устройства, симулирующие IoT
	Для различных услуг разное время выполнения: от 4 до 10 с	Dell Alienware Aurora R11 Core i7 Смоделировано 50 устройств	
PoA	TPS: 14 Задержка подтверждения транзакции: 5 мин	Нет информации	—
Raft	Для стандартного тайм-аута (150–300 мс) время определения лидера достигает 3 с	5 нод, подключенных через Ethernet-коммутатор 1 Гбит/с, со средним временем трансляции 15 мс	Основное беспокойство вызывает достижение состояния определения лидера, так как без него блокчейн-сеть не может выполнять свои функции. Производительность алгоритма при определенном лидере высока, так как транзакции выполняет только один узел, а остальные синхронизируются с ним

- 6) простота как для использования клиентами, так и для участия в качестве туманного узла;
7) поддержка автономности кластера в случае отказа узла оркестрации.

В соответствии с этими требованиями мы предлагаем архитектуру туманных вычислений с интеграцией частного блокчейна. Она также будет состоять из четырех элементов, но с некоторыми особенностями.

1. Конечное устройство (клиент). При необходимости оказания услуг клиент обращается к ближайшему туманному узлу. Кроме информации о требовании, предъявляемом к запрашиваемой услуге, клиент должен предоставить метаданные для определения приоритета выделения услуги по обработке данных (системы реального времени будут иметь приоритет выше, чем, например, системы коммунальных датчиков). Также клиент должен пройти регистрацию и получить секретный и открытый ключ. Открытый ключ имеется у всех узлов оркестрации в туманной сети.

2. Туманные узлы. В рамках данной архитектуры туманные узлы входят в некоторый кластер, который управляется одним из узлов оркестрации. Сами кластеры являются гибкими, и узлы могут как входить в них, так и выходить в зависимости от решения узла оркестрации. Узлы, кроме обработки и хранения данных, должны регулярно передавать в узел оркестрации данные о своем состоянии и информацию о своих ресурсах. Туманный узел собирает и отправляет в узел оркестрации запросы на услуги клиентов. Каждый туманный узел входит в блокчейн, однако не участвует в подтверждении транзакций из-за проблем доверия к ним и экономии ресурсов. Операции, которые назначаются узлом оркестрации, оформляются в виде смарт-контракта, который включает в себя контейнер с исполняемыми инструкциями и окружением или блоки для хранения, а также вознаграждение за оказываемую услугу. Это позволит мотивировать частных лиц для участия в туманных вычислениях.

Для непосредственного участия в туманных вычислениях узел должен пройти аутентификацию и выбрать объем ресурсов и определенные условия для их представления (например, если заряд батареи на мобильном устройстве ниже 50 %).

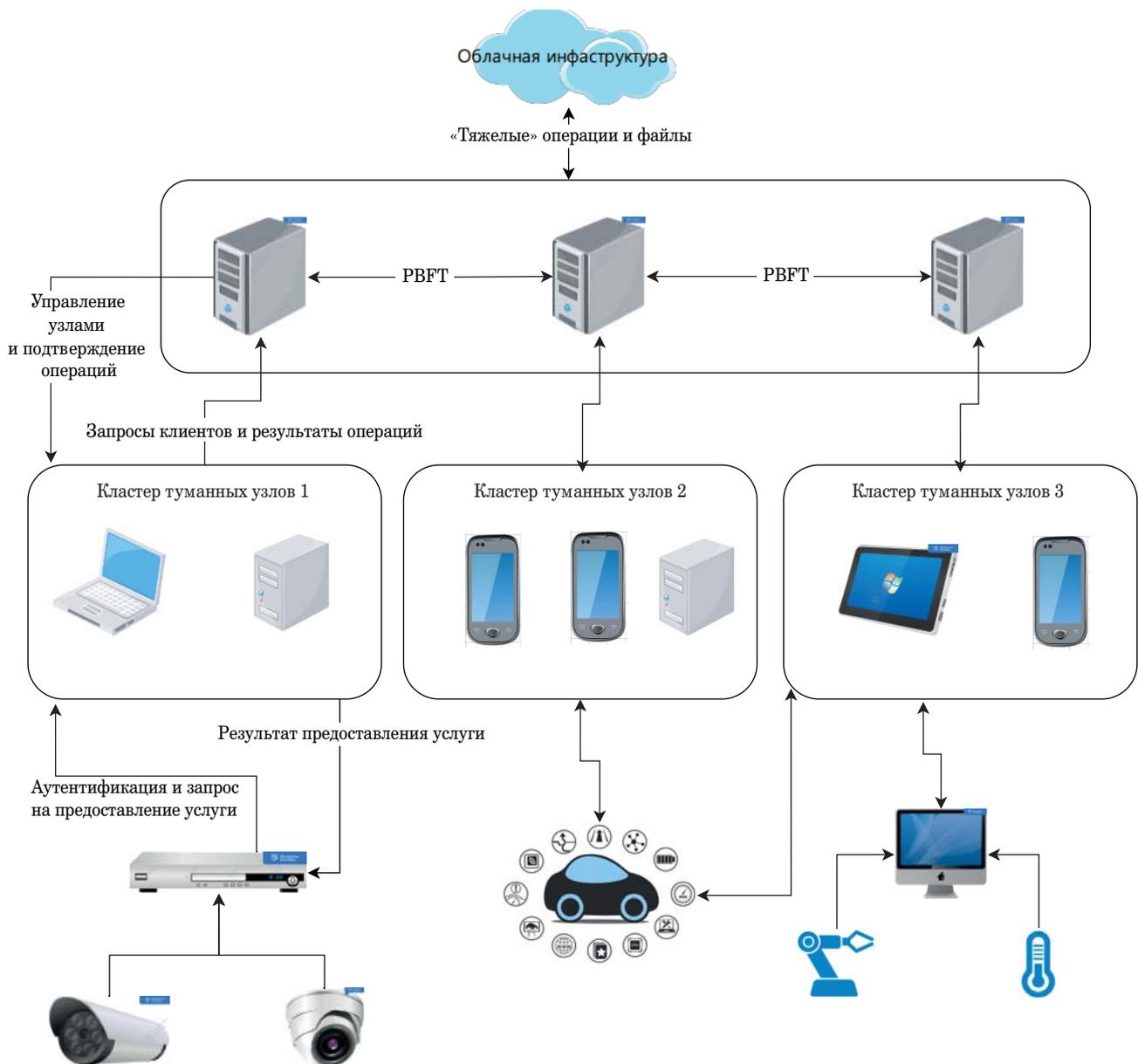
3. Узлы оркестрации. Каждый узел выполняет свои обязанности в своем кластере. Кроме того, именно эти узлы участвуют в алгоритме консенсуса и, соответственно, ведут учет бухгалтерской книги и подтверждают транзакции. Также именно эти узлы имеют связь с облачной инфраструктурой.

4. Облачная инфраструктура. Облако лежит за пределами туманных вычислений, по-

этому архитектура на него не распространяется. Однако сообщение между блокчейном и облаком выглядит также перспективным направлением для исследований [20].

На основании описанных выше требований и архитектуры можно выделить следующие критерии, которым должен удовлетворять алгоритм консенсуса:

- 1) поддержка работы алгоритма частными блокчейн-платформами;
- 2) поддержка системы мотивации участников;
- 3) возможность масштабируемости при увеличении числа узлов в сети.



■ **Рис. 2.** Общий вид архитектуры туманных вычислений с использованием технологии блокчейн
 ■ **Fig. 2.** General view of fog computing architecture using blockchain technology

Также при выборе стоит обращать внимание на скорость подтверждения транзакций и накладные расходы в случае, если алгоритм требует выбора узла лидера.

В качестве алгоритма консенсуса для узлов оркестрации выберем PBFT, так как он в большей степени соответствует заявленным критериям, и в нашей архитектуре небольшое количество узлов, которые принимают участие в консенсусе.

Алгоритм Raft тоже годится для использования, однако он требует выбора узла лидера, что нам не подходит, так как во время выборов сервер становится недоступен для оказания услуг свыше 3 с, а при сетевых задержках это время может еще увеличиться, что делает этот алгоритм неподходящим для туманных вычислений. Кроме того, Raft не поддерживается всеми блокчейн-платформами, в отличие от PBFT, который является классическим алгоритмом для приватного блокчейна.

Алгоритм PoA не подходит из-за того, что в нем невозможно задать мотивацию для участия рядовых узлов, которая крайне важна для реализации действительно качественной и расширяемой архитектуры. Кроме того, у него низкая скорость подтверждения транзакций, что также делает его применение невозможным в туманных вычислениях.

Алгоритм PoAh выглядит перспективным для использования в гибридном блокчейне, однако мы не будем рассматривать данный класс блокчейна в этой работе вследствие ограниченности по тематике и объему.

Общий вид архитектуры представлен на рис. 2.

Из преимуществ данной архитектуры можно выделить следующие принципы, поддерживаемые ею:

- безопасность;
- масштабируемость;
- открытость;
- RAS (надежность, доступность, удобство обслуживания);
- гибкость;
- иерархичность.

Литература

1. Галактионов М. А., Маколкина М. А., Киричек Р. В. Обзор протоколов сетей связи шестого поколения и сетей 2030. *СПбНТОРЭС: тр. ежегодной НТК*, 2021, № 1, с. 183–186. <https://conf-ntores.etu.ru/assets/files/2021/cp/papers/183-186.pdf> (дата обращения: 21.06.2022).
2. Плотников В. А. Цифровизация производства: теоретическая сущность и перспективы развития в российской экономике. *Известия Санкт-Петербургского государственного экономического университета*, 2018, № 4, с. 16–24.

Существуют проблемы с автономностью кластеров, так как они полностью зависимы от узла оркестрации, и отказ его работы может привести к краху существенной части туманной сети. В таком случае оставшиеся работоспособные узлы оркестрации должны временно взять на себя обязанности выбывшего узла, это также возможно за счет использования технологии блокчейн.

Одним из основных преимуществ использования блокчейна является обеспечение конфиденциальности, доступности и целостности данных в нем. Это достигается за счет использования криптографических средств, хранения хешей файлов и метаданных операций в блокчейне, а также шардирования хранимых файлов в туманной сети.

Заключение

В ходе анализа существующих предложений по архитектуре туманных вычислений выявлены проблемы, связанные с обеспечением ИБ данных, оркестрацией и неясным разделением туманных узлов, которые предоставляют услуги клиентам. Кроме того, установлено, что все разрабатываемые архитектуры не учитывают мотивацию участия частных лиц в туманной сети, что крайне важно для общественных туманных вычислений. Предлагаемая архитектура решает эти проблемы с помощью внедрения частного блокчейна, который благодаря своим свойствам позволяет безопасно пользоваться услугами туманных вычислений, а также дает мотивацию за счет вознаграждения участников туманных вычислений. Однако данная архитектура требует развертывания базовой части туманной сети и финансирования, при этом обеспечивает наибольшую жизнеспособность в рамках общественной работы и поддержания таких систем, как умная транспортная сеть и система доставки с помощью дронов.

3. Abdulkareem K. H., Mohammed M. A., Gunasekaran S. S., Almhiqani M. N., Mutlag A. A., Mostafa S. A., Ali N. S., Dheyaa A. I. A Review of fog computing and machine learning: Concepts, applications, challenges, and open issues. *IEEE Global Communications Conf. (GLOBECOM)*, 2019, pp. 1–6, doi:10.1109/ACCESS.2019.2947542
4. Khan S., Qin Y., Parkinson S. Fog computing security: A review of current applications and security solutions. *Journal of Cloud Computing*, 2017, vol. 6, no. 19, pp. 1–22. doi:10.1186/s13677-017-0090-3
5. Velasquez K., Abreu D., Assis M., Senna C., Aranha D., Bittencourt L., Laranjeiro N., Curado M., Viei-

- ra M., Monteiro E., Madeira E.** Fog orchestration for the internet of everything: State-of-the-art and research challenges. *Journal of Internet Services and Applications*, 2019, no. 9, pp. 1–23. doi:10.1186/s13174-018-0086-3
- 6. Khalid F.** Privacy and security problems in fog computing. *Communications on Applied Electronics*, 2016, no. 4, pp. 1–7. doi:10.5120/cae2016652088
- 7. Ren S., Liu B., Yang F., Wei X., Yang X., Wang C.** BlockDNS: Enhancing domain name ownership and data authenticity with blockchain. *IEEE Global Communications Conf. (GLOBECOM)*, 2019, pp. 1–6. doi:10.1109/GLOBECOM38437.2019.9013817
- 8. Conti M., Kumar G., Nerurkar P., Saha R., Vigneri L.** A survey on security challenges and solutions in the IOTA. *Journal of Network and Computer Applications*, 2022, vol. 203, pp. 1–27. doi:10.1016/j.jnca.2022.103383
- 9. Тюкалова Н. М., Разувакин А. А.** Современная концепция цифровой идентификации авиапассажира. *Научный Вестник МГТУ ГА*, 2018, т. 21, № 4, с. 39–47. doi:10.26467/2079-0619-2018-21-4-39-47
- 10. Mahmood Z., Ramachandran M.** Fog computing: Concepts, principles and related paradigms. *Fog Computing: Concepts, Frameworks and Technologies*. Springer, 2018. Pp. 3–21. doi:10.1007/978-3-319-94890-4_1
- 11. Iorga M., Feldman L., Barton R., Martin M., Goren N., Mahmoudi C.** *Fog Computing Conceptual Model*. Special Publication (NIST SP). National Institute of Standards and Technology, 2018. doi:10.6028/NIST.SP.500-325. <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.500-325.pdf> (дата обращения: 21.06.2022).
- 12. Федоров И. Р., Пименов А. В., Панин Г. А., Беззатеев С. В.** Технология блокчейн в сетях 5G: сравнение производительности частных и публичных блокчейнов. *Проблемы информационной безопасности. Компьютерные системы*, 2021, № 3(47), с. 55–62.
- 13. Pahlajani S., Kshirsagar A., Pachghare V.** Survey on private blockchain consensus algorithms. *1st Intern. Conf. on Innovations in Information and Communication Technology (ICIICT)*, 2019, pp. 1–6. doi:10.1109/ICIICT1.2019.8741353
- 14. Puthal D., Mohanty S., Yanambaka V., Koungianos E.** PoAh: A novel consensus algorithm for fast scalable private blockchain for large-scale IoT frameworks. *arXiv: CS – Cryptography and Security*, 2020, pp. 1–26. doi:10.48550/arXiv.2001.07297
- 15. Dongyan H., Xiaoli M., Shengli Z.** Performance analysis of the Raft consensus algorithm for private blockchains. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2020, vol. 50, iss. 1, pp. 172–181. doi:10.1109/TSMC.2019.2895471
- 16. Saide Z., Zhipeng C., HuaFu H., Yingshu L., Wei L.** zkCrowd: A hybrid blockchain-based crowdsourcing platform. *IEEE Transactions on Industrial Informatics*, 2020, vol. 16, iss. 6, pp. 4196–4205. doi:10.1109/TII.2019.2941735
- 17. Latif S., Idrees Z., Ahmad J., Lirong Z., Zou Z.** A blockchain-based architecture for secure and trustworthy operations in the industrial Internet of Things. *Journal of Industrial Information Integration*, 2021, vol. 21. doi:10.1016/j.jii.2020.100190
- 18. Kaur M., Khan M. Z., Gupta S., Noorwali A., Chakraborty C., and Pani S. K.** MBSP: Performance analysis of large scale mainstream blockchain consensus protocols. *IEEE Access*, 2021, vol. 9, pp. 80931–80944. doi:10.1109/ACCESS.2021.3085187
- 19. Schäffer M., Di Angelo M., Salzer G.** Performance and scalability of private ethereum blockchains. *Business Process Management: Blockchain and Central and Eastern Europe Forum. Lecture Notes in Business Information Processing*. Springer, 2019. No. 361. Pp. 103–118. doi:10.1007/978-3-030-30429-4_8
- 20. Беззатеев С. В., Федоров И. Р.** Технология блокчейн в сетях 5G. *Научно-технический вестник информационных технологий, механики и оптики*, 2020, т. 20, № 4(128), с. 472–484. doi:10.17586/2226-1494-2020-20-4-472-484

UDC 004.56

doi:10.31799/1684-8853-2022-5-40-48

EDN: KJPXLT

Designing fog computing architecture with the use of blockchain technologyA. V. Pimenov^a, Student, orcid.org/0000-0002-9136-3514I. R. Fedorov^a, Post-Graduate Student, orcid.org/0000-0003-2422-4714S. V. Bezzateev^b, Dr. Sc., Tech, Professor, orcid.org/0000-0002-0924-6221, bezzateev_sergey@mail.ru^aITMO University, 49, Kronverksky Pr., 197101, Saint-Petersburg, Russian Federation^bSaint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaya St., 190000, Saint-Petersburg, Russian Federation**Introduction:** Due to the growth in the number and variety of devices connected to the Internet, the requirements for network performance and data transmission security are increasing. Today, performance problems are usually solved through cloud, fog and

edge computing, while the problem of data storage and transmission security remains relevant. One of the effective ways to solve this problem is to use blockchain technology. **Purpose:** Designing the architecture of a fog computing network based on blockchain technology. **Results:** Based on the research in the field of fog computing, the requirements for the fog computing architecture were determined, such as: autonomy, scalability, flexibility, hierarchy, security, reliability, availability, serviceability. The selected criteria for building an architecture led to the choice in favor of a private blockchain due to its higher performance compared to a public blockchain. A comparative analysis of the consensus algorithms that are most often used in private blockchains was carried out and the most suitable one was chosen. Based on the requirements put forward and the results of the analysis, a fog computing architecture model based on a private blockchain was designed. The architecture consists of four elements: end devices, fog nodes, orchestration nodes, and cloud infrastructure. The blockchain includes fog nodes and orchestration nodes, which ensures the confidentiality, availability and integrity of data in the fog network. **Practical relevance:** Paper results can be used in the design of fog computing networks both separately and as part of 5G mobile networks.

Keywords – fog computing, fog computing architecture, blockchain, information security, orchestration, Internet of Things.

For citation: Pimenov A. V., Fedorov I. R., Bezzateev S. V. Designing fog computing architecture with the use of blockchain technology. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 5, pp. 40–48 (In Russian). doi:10.31799/1684-8853-2022-5-40-48, EDN: KJPXLT

References

- Galaktionov M. A., Makolkina M. A., Kirichek R. V. Overview of the protocols of the sixth generation communication networks and networks 2030. *SPbNTORES: tr. ezhegodnoj NTK* [Proc. 76th Scientific and Technical Conference of St. Petersburg NTO RES im. A. S. Popova, Dedicated to the Radio Day]. Saint-Petersburg, 2021, no. 1, pp. 183–186. Available at: <https://conf-ntores.etu.ru/assets/files/2021/cp/papers/183-186.pdf> (accessed 21 June 2022) (In Russian).
- Plotnikov V. A. Digitalization of production: the theoretical essence and development prospects in the Russian economy. *Izvestiya Sankt-Peterburgskogo gosudarstvennogo ekonomicheskogo universiteta*, 2018, no. 4, pp. 16–24 (In Russian).
- Abdulkareem K. H., Mohammed M. A., Gunasekaran S. S., Almhqani M. N., Mutlag A. A., Mostafa S. A., Ali N. S., Dheyaa A. I. A review of fog computing and machine learning: Concepts, applications, challenges, and open issues. *IEEE Global Communications Conf. (GLOBECOM)*, 2019, pp. 1–6. doi:10.1109/ACCESS.2019.2947542
- Khan S., Qin Y., Parkinson S. Fog computing security: A review of current applications and security solutions. *Journal of Cloud Computing*, 2017, vol. 6, no. 19, pp. 1–22. doi:10.1186/s13677-017-0090-3
- Velasquez K., Abreu D., Assis M., Senna C., Aranha D., Bitencourt L., Laranjeiro N., Curado M., Vieira M., Monteiro E., Madeira E. Fog orchestration for the internet of everything: State-of-the-art and research challenges. *Journal of Internet Services and Applications*, 2019, no. 9, pp. 1–23. doi:10.1186/s13174-018-0086-3
- Khalid F. Privacy and security problems in fog computing. *Communications on Applied Electronics*, 2016, no. 4, pp. 1–7. doi:10.5120/cae2016652088
- Ren S., Liu B., Yang F., Wei X., Yang X., Wang C. BlockDNS: Enhancing domain name ownership and data authenticity with blockchain. *IEEE Global Communications Conf. (GLOBECOM)*, 2019, pp. 1–6. doi:10.1109/GLOBECOM.38437.2019.9013817
- Conti M., Kumar G., Nerurkar P., Saha R., Vigneri L. A survey on security challenges and solutions in the IOTA. *Journal of Network and Computer Applications*, 2022, vol. 203, pp. 1–27. doi:10.1016/j.jnca.2022.103383
- Tyukalova N. M., Razuvaikin A. A. Modern concept of digital identification of air passengers. *Civil Aviation High Technologies*, 2018, vol. 21, no. 4, pp. 39–47 (In Russian). doi:10.26467/2079-0619-2018-21-4-39-47
- Mahmood Z., Ramachandran M. *Fog computing: Concepts, principles and related paradigms*. In: *Fog Computing: Concepts, Frameworks and Technologies*. Springer, 2018. Pp. 3–21. doi:10.1007/978-3-319-94890-4_1
- Iorga M., Feldman L., Barton R., Martin M., Goren N., Mahmoudi C. *Fog Computing Conceptual Model*. Special Publication (NIST SP). National Institute of Standards and Technology, 2018. doi:10.6028/NIST.SP.500-325. Available at: <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.500-325.pdf> (accessed 21 June 2022).
- Fedorov I. R., Pimenov A. V., Pamim G. A., Bezzateev S. V. Blockchain in 5G networks: Performance comparison of private and public blockchain. *Information Security Problems. Computer Systems*, 2021, no. 3(47), pp. 55–62 (In Russian).
- Pahlajani S., Kshirsagar A., Pachghare V. Survey on private blockchain consensus algorithms. *1st Intern. Conf. on Innovations in Information and Communication Technology (ICIICT)*, 2019, pp. 1–6. doi:10.1109/ICIICT1.2019.8741353
- Puthal D., Mohanty S., Yanambaka V., Kougianos E. PoAh: A novel consensus algorithm for fast scalable private blockchain for large-scale IoT frameworks. *arXiv: CS – Cryptography and Security*, 2020, pp. 1–26. doi:10.48550/arXiv.2001.07297
- Dongyan H., Xiaoli M., Shengli Z. Performance analysis of the Raft consensus algorithm for private blockchains. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2020, vol. 50, iss. 1, pp. 172–181. doi:10.1109/TSMC.2019.2895471
- Saide Z., Zhipeng C., Huafu H., Yingshu L., Wei L. zkCrowd: A hybrid blockchain-based crowdsourcing platform. *IEEE Transactions on Industrial Informatics*, 2020, vol. 16, iss. 6, pp. 4196–4205. doi:10.1109/TII.2019.2941735
- Latif S., Idrees Z., Ahmad J., Lirong Z., Zou Z. A blockchain-based architecture for secure and trustworthy operations in the industrial Internet of Things. *Journal of Industrial Information Integration*, 2021, vol. 21. doi:10.1016/j.jii.2020.100190
- Kaur M., Khan M. Z., Gupta S., Noorwali A., Chakraborty C., and Pani S. K. MBP: Performance analysis of large scale mainstream blockchain consensus protocols. *IEEE Access*, 2021, vol. 9, pp. 80931–80944. doi:10.1109/ACCESS.2021.3085187
- Schäffer M., Di Angelo M., Salzer G. *Performance and scalability of private ethereum blockchains*. In: *Business Process Management: Blockchain and Central and Eastern Europe Forum. Lecture Notes in Business Information Processing*. Springer, 2019. Vol. 361. Pp. 103–118. doi:10.1007/978-3-030-30429-4_8
- Bezzateev S. V., Fedorov I. R. Blockchain technology in 5G networks. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2020, vol. 20, no. 4(128), pp. 472–484 (In Russian). doi:10.17586/2226-1494-2020-20-4-472-484



Analyzing and stabilizing multichannel ALOHA with the use of the preamble-based exploration phase

A. A. Burkov^a, Assistant Professor, orcid.org/0000-0002-0920-585X

R. O. Rachugin^a, Master Student, orcid.org/0000-0001-5813-3867

A. M. Turlikov^a, Dr. Sc., Tech., Professor, orcid.org/0000-0001-7132-094X, turlikov@vu.spb.ru

^aSaint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaya St., 190000, Saint-Petersburg, Russian Federation

Introduction: Internet of Things devices are actively used within the framework of Massive Machine-Type Communication scenarios. The interaction of devices is carried out by random multiple-access algorithms with limited throughput. To improve throughput one can use orthogonal preambles in the ALOHA-type class of algorithms. **Purpose:** To analyze ALOHA-based algorithms using the exploration phase and to calculate the characteristics for the algorithm with and without losses with a finite number of channels. **Results:** We have described a system model that employs random access for data transmission over a common communication channel with the use of orthogonal preambles and exploration phase. We have obtained a formula for numerical calculation of the throughput of an algorithm channel with losses with an infinite number of preambles and a given finite number of channels. The calculation results for several values of the number of independent channels are presented. A modification of the algorithm using the exploration phase and repeated transmissions is proposed and described. The system in question can work without losses. For this system, we have given the analysis of the maximum input throughput up to which the system operates stably. Also, the average delay values for the algorithm that were obtained by simulation modeling are shown. By reducing the number of available preambles, the results obtained can be used as an upper bound on the system throughput. **Practical relevance:** The results obtained allow to assess the potential for improving the throughput of random multiple-access systems in 6G networks through the application of the exploration phase.

Keywords – multichannel ALOHA, maximum throughput, random multiple access, stability, Massive Machine-Type Communications, Internet of Things.

For citation: Burkov A. A., Rachugin R. O., Turlikov A. M. Analyzing and stabilizing multichannel ALOHA with the use of the preamble-based exploration phase. *Informatsionno-upravlyaiushchie sistemy* [Information and Control Systems], 2022, no. 5, pp. 49–59. doi:10.31799/1684-8853-2022-5-49-59, EDN: KGVYDI

Introduction

Currently, there is an active development of wireless communication technologies, caused by a large increase in the requirements for the number of operating devices and the amount of transmitted data. A prime example is the Internet of Things (IoT) technology. Support for this technology in cellular networks is being studied in both the standard (5G) and emerging standards (6G). IoT systems are expected to work within the framework of the Massive Machine-Type Communications scenario. The possibility of operation of a potentially infinite number of devices and a high load on the network [1–3] are considered. Since the number of users is potentially infinite, random access algorithms are assumed [4, 5]. As a rule, varieties of algorithms such as ALOHA and its modifications are considered. The existing algorithms have restrictions on the limiting input arrival rate (throughput), which become few with the existing increase in the number of users. There are various approaches to increase the throughput of algorithms. Thus, approaches

based on the use of non-orthogonal multiple access (NOMA), sequential interference cancellation (SIC) and orthogonal preambles are proposed [6–12]. However, NOMA and SIC methods are computationally intensive. Therefore, in this article we will consider an approach using orthogonal preambles. The work [12] considers a system with independent channels and a multichannel ALOHA algorithm that uses orthogonal preambles to implement work with two phases. The exploration phase is the phase of evaluating possible collisions using preambles. Users then transmit messages in the Data Transmission Phase based on the results of the EP. In this paper, we will take as a basis the model of such a system. In contrast to [12], we will analyze the throughput for a finite number of independent channels. It also proposes a modification of the algorithm using repeated transmissions, which allows the system to work without losses. This algorithm can be used to organize data transmission in high reliability systems [13–15]. For such a model, the limiting input arrival rate is sought, up to which the system is stable for a given number of channels.

Analysis of a system with losses

Let us introduce a system of assumptions for the system under consideration.

Assumption 1. The entire time of the channel operation is divided into frames of the same length. A frame consists of two phases: an exploration phase (EP) and a data transmission phase (DTP). Each phase is followed by a response from the base station (BS). In the exploration phase, orthogonal preambles are transmitted, and in the data transmission phase, user messages are transmitted. Users know exactly the boundaries of the division into frames and phases and can transmit data only at the beginning of the corresponding phase at the beginning of a new frame.

Assumption 2. There is a set of K independent channels. All users have unique preambles. A preamble is a sequence of bits of a certain length, which is much less than the length of the message. The preamble transmission rate is the same. The BS can reliably determine the number of preambles on each channel.

Assumption 3. In each channel, during each phase of the frame, one of the following events can occur:

- “Success” – if the data was transmitted by one user;
- “Empty” – none of the users transmitted data;
- “Conflict” – simultaneously, two or more users were transmitting data. If a conflict occurs during the EP, then the BS will successfully determine the number of preambles. If a conflict occurs during the DTP, then the users’ messages overlap and cannot be processed on the receiving side.

At the end of each phase, all users will reliably know what event happened in the channel.

Assumption 4. The system has a potentially unlimited number of users. The number of messages appearing in the system in one frame is distributed according to the Poisson law with the parameter $K\lambda$.

Assumption 5. The time of the data transmission phase will be taken as a unit of time, and the time of the exploration phase will be taken as 0.

Let us describe the considered algorithm. According to *Assumption 1*, each frame consists of two phases: EP and DTP. In EP, users randomly select one of the K channels and send a preamble over it. The BS then estimates the number of preambles in each channel and broadcasts this information to all users. After that, users are divided into two groups. The first group (G_S) includes users in whose channels one preamble was transmitted. All other users fall into the second group (G_C). Channels selected by users in G_S are assigned to them, and all other channels belong to the second group.

Then comes DTP phase. G_S users transmit data with a probability of one. Users of the G_C decide whether to transmit data to them with a probability

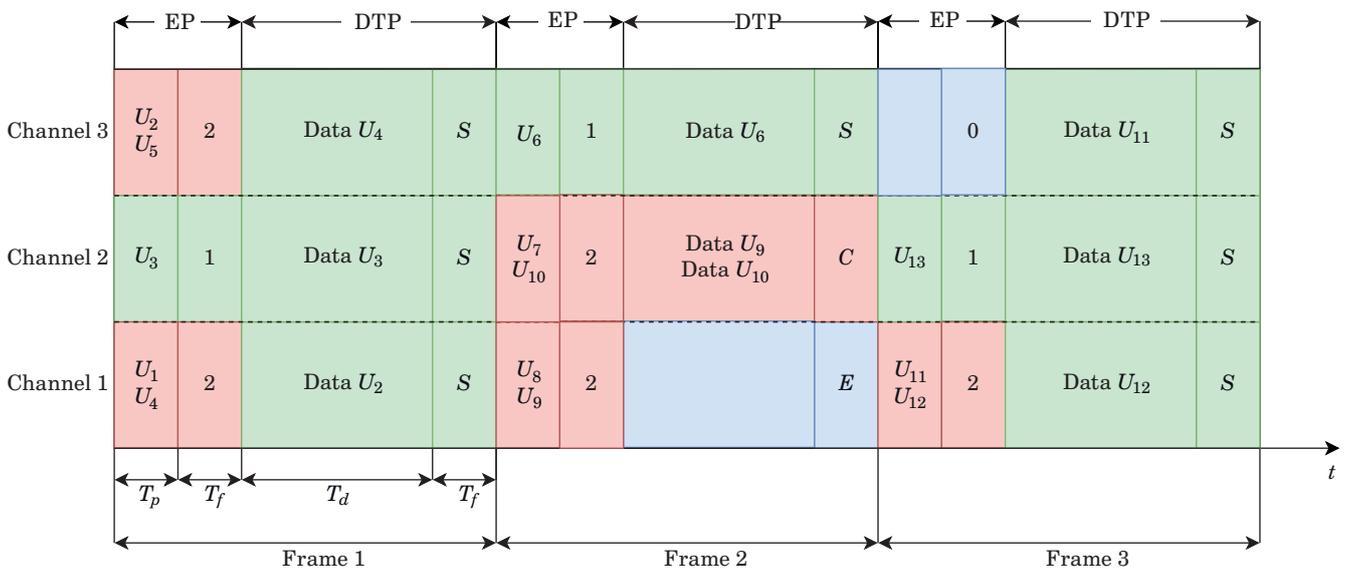
$$\text{equal to } p_{DTP} = \min\left\{1, \frac{K - |G_S|}{|G_C|}\right\}.$$

Users of the

G_C group who decide to transmit data randomly select one of the channels of the second group and transmit their data. The base station then informs all users about the events that have occurred in each channel via the feedback channel.

Consider an example of the algorithm shown in Fig. 1.

The system has three data channels. Consider the first frame. In the first channel in the exploration



■ **Fig. 1.** Example of the algorithm’s lossy work

phase users U_1 and U_4 decided to send their preambles, in the second channel user U_3 decided to send a preamble, in the third channel users U_2 and U_5 decided to send preambles. After evaluating the preambles in each channel, the base station will return to all users the next set of data $\{2, 1, 2\}$. Thus, channels one and three belong to group two, hence users U_1, U_2, U_4 and U_5 also belong to the G_C . Channel two belongs to group one and also user U_3 belongs to G_S .

This is followed by the data transmission phase, in which the user from the first group remains in the same channel and transmits with probability equal to one. Users of the second group randomly decide whether or not to send data to them. In this case, the transmission probability will be equal to

$$p_{DTP} = \min\left\{1, \frac{2}{4}\right\} = 0.5. \text{ Assume that users } U_1$$

and U_5 decide not to transmit, but users U_2 and U_4 decide to transmit and randomly reselect one of the channels of group two. Let user U_2 reselect the first channel, and user U_4 reselect the third channel. Thus, as a result, the "Success" event will occur in all channels.

Consider the next frame. Suppose that in the exploration phase, users U_8 and U_9 decide to transmit their preambles on the first channel, users U_7 and U_{10} on the second channel, and user U_6 on the third channel. Then, the base station, having estimated the number of received preambles in each channel, returned the following data set to all users $\{2, 2, 1\}$. Thus, channels one and two belong to group two, hence users U_7, U_8, U_9 and U_{10} also belong to the G_C . Channel three belongs to group one, and user U_6 belongs to G_S .

This is followed by the data transmission phase, in which the user from the first group remains in the same channel and transmits with probability equal to one. Users of the second group randomly decide whether or not to send data to them. In this case, the probability of transmission will be equal

$$\text{to } p_{DTP} = \min\left\{1, \frac{2}{4}\right\} = 0.5. \text{ Assume that users } U_7$$

and U_8 decide not to transmit, users U_9 and U_{10} decide to transmit and randomly reselect one of the channels of group two. Let both users U_9 and U_{10} reselect the second channel. Thus, the event "Empty" occurred in the first channel, the event "Conflict" occurred in the second channel, and the event "Success" occurred in the third channel.

Consider the third frame. Assume that in the exploration phase, users U_{11} and U_{12} decide to transmit preambles on the first channel, user U_{13} transmits the preamble on the second channel. Then, the base station, having estimated the number of received preambles in each channel, returned to all

users the following data set $\{2, 1, 0\}$. Thus, channels one and three belong to group two, hence users U_{11} and U_{12} belong to the G_C . Channel two belongs to group one and also user U_{13} belongs to G_S .

This is followed by the data transmission phase, in which the user from the first group remains in the same channel and transmits with probability equal to one. The users of group two randomly decide to send data to them or not. In this case, the probability of transmission will be equal to

$$p_{DTP} = \min\left\{1, \frac{2}{2}\right\} = 1. \text{ Assume that user } U_{11} \text{ de-}$$

cidies to transmit and randomly reselects the third channel. User U_{12} decided to transmit and randomly reselected the first channel. Thus, the "Success" event occurred in all channels.

Let us analyze for the algorithm with a finite number of channels. By analogy with [12], we will consider the normalized throughput (per channel) and denote it as $T(\lambda, K)$.

Throughput per channel is calculated as

$$T(\lambda, K) = \frac{E[N_t]}{K}, \quad (1)$$

where N_t is the number of users who logged out of the system at time t , $N_t \in \{0, 1, 2, \dots, K\}$; K is the number of channels in the system; λ is the intensity of the input arrival rate in the system, which affects the distribution of the random variable N_t .

Denote by N_t^i the number of users leaving channel I at time t , then $N_t = N_t^{(1)} + N_t^{(2)} + \dots + N_t^{(K)}$. Then from (1) it follows:

$$T(\lambda) = \frac{E[N_t^{(1)} + N_t^{(2)} + \dots + N_t^{(K)}]}{K} = \frac{\sum_{i=1}^K E[N_t^{(i)}]}{K}.$$

It might be noted that $E[N_t^{(1)}] = E[N_t^{(2)}] = \dots = E[N_t^{(K)}]$. Then:

$$T(\lambda) = E[N_t^{(1)}].$$

Taking into account the fact that only message one user can be successfully transmitted in one channel, then $E[N_t^{(1)}] = \Pr\{N_t^{(1)} = 1\}$. Thus, we get that:

$$T(\lambda) = \Pr\{N_t^{(1)} = 1\}.$$

Next, we will show how to calculate this probability for channel number 1. For a channel with other numbers, the probability is calculated similarly. Note the following:

$$\Pr\{N_t^{(1)} = 1\} = \Pr\{N_t^{(1)} = 1, S_{EP}\} + \Pr\{N_t^{(1)} = 1, S_{DTP}\},$$

where S_{EP} is success in the first phase (EP); S_{DTP} is success in the second phase (DTP).

Success in the first phase occurs only if the channel in question has been selected by one user. Given *Assumption 4*, the value of this probability can be calculated as

$$\Pr\{N_t^{(1)} = 1, S_{EP}\} = \lambda e^{-\lambda}.$$

If there was no success in a given channel in the first phase, then success in the second phase depends on the number of channels in which there was no success, let's denote the number of such channels as L , then:

$$\Pr\{N_t^{(1)} = 1, S_{DTP}\} = \sum_{l=1}^K \Pr\{N_t^{(1)} = 1, S_{DTP}, L = l\}. \quad (2)$$

Next, we will show how to calculate the terms from the sum in expression (2), for each l . Denote by $\Pr\{S_{EP}, W = K - l\}$ the probability that the event "Success" occurred in $K - l$ channels during the exploration phase.

If $l = 1$, then in all other channels except the current one there should be a "Success" event, therefore

$$\Pr\{S_{EP}, W = K - 1\} = (\lambda e^{-\lambda})^{K-1}.$$

Let i_1 users arrive at the input of the current channel. We write the probability of such an event as: $\Pr\{\text{entered } i_1 \text{ users}\} = \frac{\lambda^{i_1}}{i_1!} e^{-\lambda}$. During the data transmission phase, the current channel can have a "Success" event if two events occur:

- at the exploration phase of the current channel there was an event "Conflict" ($i_1 \geq 2$);
- during the data transmission phase, only one of the users participating in the "Conflict" event during the exploration phase will make a decision to transmit in the current channel. Since in this case, the users make a decision to transmission with the probability $\frac{1}{i_1}$. So, the probability of such an event is

$$\Pr\{S_{DTP} | i_1\} = C_{i_1}^1 \frac{1}{i_1} \left(1 - \frac{1}{i_1}\right)^{i_1-1} = \left(1 - \frac{1}{i_1}\right)^{i_1-1}.$$

Taking into account the introduced notation, we obtain:

$$\begin{aligned} \Pr\{N = 1, S_{DTP}, L = 1\} &= \Pr\{S_{EP}, W = K - 1\} \sum_{i_1=2}^{\infty} \Pr\{S_{DTP} | i_1\} \Pr\{\text{entered } i_1 \text{ users}\} = \\ &= \lambda^{K-1} e^{-\lambda K} \sum_{i_1=2}^{\infty} \left(1 - \frac{1}{i_1}\right)^{i_1-1} \frac{\lambda^{i_1}}{i_1!}. \end{aligned}$$

If $l = 2$, then in all channels, except for the current and one more, there should be a "Success" event, therefore

$$\Pr\{S_{EP}, W = K - 2\} = C_{K-1}^{K-2} (\lambda e^{-\lambda})^{K-2}.$$

Let i_1 users arrive at the input of the current channel, and i_2 users at the input of another. We write the probability of such an event as:

$$\Pr\{\text{entered } i_1 \text{ users}\} \Pr\{\text{entered } i_2 \text{ users}\} = \frac{\lambda^{i_1}}{i_1!} e^{-\lambda} \frac{\lambda^{i_2}}{i_2!} e^{-\lambda}.$$

During the data transmission phase, the current channel can have a "Success" event if two events occur:

- $i_1 \neq 1$ and $i_2 \neq 1$ and $i_1 + i_2 \geq 2$;

– during the data transmission phase, only one of the users participating in the “Conflict” event during the exploration phase will make a decision to transmit in the current channel. Since in this case the users decide to transmit in the current channel with the probability $\frac{1}{2} \cdot \frac{2}{i_1 + i_2}$. Then, the probability of the event “Success”

in the current channel is equal to $\Pr\{S_{DTP} | i_1, i_2\} = C_{i_1+i_2}^1 \frac{1}{i_1 + i_2} \left(1 - \frac{1}{i_1 + i_2}\right)^{i_1+i_2-1} = \left(1 - \frac{1}{i_1 + i_2}\right)^{i_1+i_2-1}$.

Taking into account the introduced notation, we obtain:

$$\begin{aligned} \Pr\{N = 1, S_{DTP}, L = 2\} &= C_{K-1}^{K-2} (\lambda e^{-\lambda})^{K-2} \sum_{\substack{i_1=0 \\ i_1 \neq 1 \\ i_1+i_2 \neq 0}}^{\infty} \sum_{\substack{i_2=0 \\ i_2 \neq 1 \\ i_1+i_2 \neq 0}}^{\infty} \left(1 - \frac{1}{i_1 + i_2}\right)^{i_1+i_2-1} \frac{\lambda^{i_1}}{i_1!} e^{-\lambda} \frac{\lambda^{i_2}}{i_2!} e^{-\lambda} = \\ &= C_{K-1}^1 (\lambda e^{-\lambda})^{K-2} \sum_{\substack{i_1=0 \\ i_1 \neq 1 \\ i_1+i_2 \neq 0}}^{\infty} \sum_{\substack{i_2=0 \\ i_2 \neq 1 \\ i_1+i_2 \neq 0}}^{\infty} \left(1 - \frac{1}{i_1 + i_2}\right)^{i_1+i_2-1} \frac{\lambda^{i_1}}{i_1!} e^{-\lambda} \frac{\lambda^{i_2}}{i_2!} e^{-\lambda} = \\ &= C_{K-1}^1 \lambda^{K-2} e^{-\lambda K} \sum_{\substack{i_1=0 \\ i_1 \neq 1 \\ i_1+i_2 \neq 0}}^{\infty} \sum_{\substack{i_2=0 \\ i_2 \neq 1 \\ i_1+i_2 \neq 0}}^{\infty} \left(1 - \frac{1}{i_1 + i_2}\right)^{i_1+i_2-1} \frac{\lambda^{i_1+i_2}}{i_1! i_2!}. \end{aligned}$$

If $l = 3$, then in all channels, except for the current one and two more, there should be a “Success” event, therefore

$$\Pr\{S_{EP}, W = K - 3\} = C_{K-1}^{K-3} (\lambda e^{-\lambda})^{K-3}.$$

Let i_1 users arrive at the input of the current channel, and i_2 and i_3 users, respectively, at the input of the other two channels. We write the probability of such an event as:

$$\Pr\{\text{entered } i_1 \text{ users}\} \Pr\{\text{entered } i_2 \text{ users}\} \Pr\{\text{entered } i_3 \text{ users}\} = \frac{\lambda^{i_1}}{i_1!} e^{-\lambda} \frac{\lambda^{i_2}}{i_2!} e^{-\lambda} \frac{\lambda^{i_3}}{i_3!} e^{-\lambda}.$$

During the data transmission phase, the current channel can have a “Success” event if two events occur:

– $i_1 \neq 1$ and $i_2 \neq 1$ and $i_3 \neq 1$ and $i_1 + i_2 + i_3 \geq 2$;

– during the data transmission phase, only one of the users participating in the “Conflict” event during the exploration phase will make a decision to transmit in the current channel. Since in this case the users decide to transmit in the current channel, the probability depends on the sum of the number of users in the channels under consideration. If $2 \leq i_1 + i_2 + i_3 \leq 3$, then the probability of the “Success” event in the current

channel is $\Pr\{S_{DTP} | i_1, i_2, i_3, L = 3\} = (i_1 + i_2 + i_3) \frac{1}{3} \left(1 - \frac{1}{3}\right)^{i_1+i_2+i_3-1}$. If $i_1 + i_2 + i_3 > 3$, then the probability of

the “Success” event in the current channel is $\Pr\{S_{DTP} | i_1, i_2, i_3, L = 3\} = \left(1 - \frac{1}{i_1 + i_2 + i_3}\right)^{i_1+i_2+i_3-1}$.

We introduce the following indicator function $I\{\text{statement}\} = \begin{cases} 1, & \text{true} \\ 0, & \text{false} \end{cases}$.

Taking into account the introduced notation, we obtain:

$$\begin{aligned} \Pr\{N = 1, S_{DTP}, L = 3\} &= \\ &= C_{K-1}^{K-2} (\lambda e^{-\lambda})^{K-3} \sum_{\substack{i_1=0 \\ i_1 \neq 1 \\ i_1+i_2+i_3 \neq 0}}^{\infty} \sum_{\substack{i_2=0 \\ i_2 \neq 1 \\ i_1+i_2+i_3 \neq 0}}^{\infty} \sum_{\substack{i_3=0 \\ i_3 \neq 1 \\ i_1+i_2+i_3 \neq 0}}^{\infty} \Pr\{S_{DTP} | i_1, i_2, i_3, L = 3\} \frac{\lambda^{i_1}}{i_1!} e^{-\lambda} \frac{\lambda^{i_2}}{i_2!} e^{-\lambda} \frac{\lambda^{i_3}}{i_3!} e^{-\lambda} = \end{aligned}$$

$$\begin{aligned}
 &= C_{K-1}^2 (\lambda e^{-\lambda})^{K-3} \sum_{\substack{i_1=0 \\ i_1 \neq 1 \\ i_1+i_2+i_3 \neq 0}}^{\infty} \sum_{\substack{i_2=0 \\ i_2 \neq 1 \\ i_1+i_2+i_3 \neq 0}}^{\infty} \sum_{\substack{i_3=0 \\ i_3 \neq 1 \\ i_1+i_2+i_3 \neq 0}}^{\infty} \Pr\{S_{DTP} | i_1, i_2, i_3, L=3\} \frac{\lambda^{i_1}}{i_1!} e^{-\lambda} \frac{\lambda^{i_2}}{i_2!} e^{-\lambda} \frac{\lambda^{i_3}}{i_3!} e^{-\lambda} = \\
 &= C_{K-1}^2 \lambda^{K-3} e^{-\lambda K} \sum_{\substack{i_1=0 \\ i_1 \neq 1 \\ i_1+i_2+i_3 \neq 0}}^{\infty} \sum_{\substack{i_2=0 \\ i_2 \neq 1 \\ i_1+i_2+i_3 \neq 0}}^{\infty} \sum_{\substack{i_3=0 \\ i_3 \neq 1 \\ i_1+i_2+i_3 \neq 0}}^{\infty} \Pr\{S_{DTP} | i_1, i_2, i_3, L=3\} \frac{\lambda^{i_1+i_2+i_3}}{i_1! i_2! i_3!} = \\
 &= C_{K-1}^2 \lambda^{K-3} e^{-\lambda K} \sum_{\substack{i_1=0 \\ i_1 \neq 1 \\ i_1+i_2+i_3 \neq 0}}^{\infty} \sum_{\substack{i_2=0 \\ i_2 \neq 1 \\ i_1+i_2+i_3 \neq 0}}^{\infty} \sum_{\substack{i_3=0 \\ i_3 \neq 1 \\ i_1+i_2+i_3 \neq 0}}^{\infty} \left(I\{i_1+i_2+i_3 \leq 3\} \left((i_1+i_2+i_3) \frac{1}{3} \left(1-\frac{1}{3}\right)^{i_1+i_2+i_3-1} \right) + \right. \\
 &\quad \left. + I\{i_1+i_2+i_3 > 3\} \left(\left(1-\frac{1}{i_1+i_2+i_3}\right)^{i_1+i_2+i_3-1} \right) \right) \frac{\lambda^{i_1+i_2+i_3}}{i_1! i_2! i_3!}.
 \end{aligned}$$

Then we can write in general:

$$\begin{aligned}
 T(\lambda) = \lambda e^{-\lambda} + \sum_{l=1}^K C_{K-1}^{l-1} \lambda^{K-l} e^{-\lambda K} \sum_{\substack{i_1=0 \\ i_1 \neq 1 \\ \sum_{j=1}^l i_j \neq 0}}^{\infty} \dots \sum_{\substack{i_l=0 \\ i_l \neq 1 \\ \sum_{j=1}^l i_j \neq 0}}^{\infty} \left(I\left\{ \sum_{j=1}^l i_j \leq l \right\} \sum_{j=1}^l i_j \frac{1}{l} \left(1-\frac{1}{l}\right)^{\sum_{j=1}^l i_j} + \right. \\
 \left. + I\left\{ \sum_{j=1}^l i_j > l \right\} \left(1-\frac{1}{\sum_{j=1}^l i_j}\right)^{\sum_{j=1}^l i_j-1} \right) \frac{\lambda^{\sum_{j=1}^l i_j}}{\prod_{j=1}^l i_j!}, \tag{3}
 \end{aligned}$$

where $I\{statement\} = \begin{cases} 1, & \text{true} \\ 0, & \text{false} \end{cases}$.

Formula (3) can be used to calculate for any number of channels.

Figure (2) shows the results obtained by (3) for a different number of channels depending on the throughput per channel.

From Fig. 2, it can be seen that the throughput limit is reached when the input arrival rate is greater than 1. The values of the maximum throughput and the corresponding input arrival rate are presented in Table 1.

The values presented in this table will be used as a parameter for the lossless algorithm.

Analysis of a system without losses

Let us describe a modification of the algorithm [12], in which users remain in the system until they successfully transmit a message (system without losses).

To do this, we add one more assumption to the system.

Assumption 6. The number of active users (having a message ready for transmission) is known to M .

According to *Assumption 1*, each frame consists of two phases: EP and DTP. This algorithm repeats the description structure of the phases of the previous algorithm, but the transmission probability in the EP phase is

$p_{EP} = \min\left\{1, \frac{KG}{M}\right\}$, where G is the algorithm parameter. The user tries to send his message until he receives confirmation of successful transmission.

Consider an example of the algorithm shown in Fig. 3.

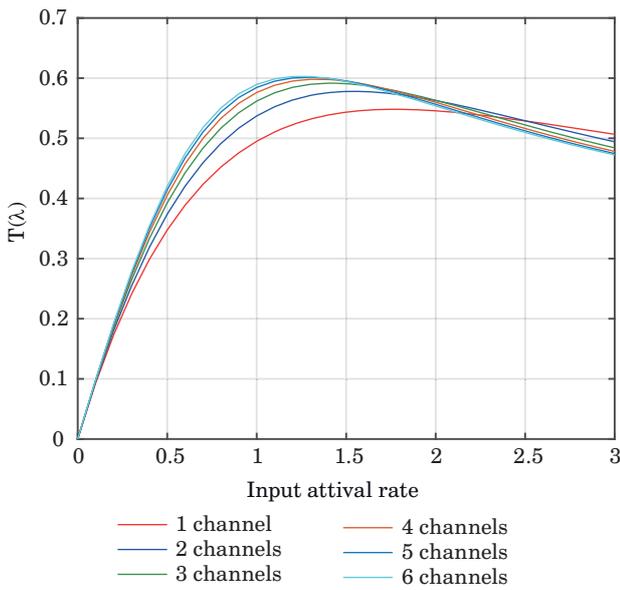


Fig. 2. Throughput per channel on the input arrival rate

Consider the first frame. Suppose that by the beginning of this frame there are messages from users U_1, U_2, U_3, U_4, U_5 and everyone decided to transmit. Let in the EP users U_4 and U_5 decide to transmit in the first channel, users U_3 in the second channel, users U_1 and U_2 in the third channel. BS, having estimated the number of preambles in each channel, will return the following information to all users

Table 1. Maximum throughput

Number channel	max $T(\lambda)$	λ
1	0.5482	1.775
2	0.578	1.55
3	0.5917	1.415
4	0.5983	1.34
5	0.6015	1.285
6	0.6031	1.25

{2, 1, 2}. Thus, user U_3 belongs to the G_S , and users U_1, U_2, U_4, U_5 belong to the G_C .

Then the DTP follows, in which the U_3 user transmits with probability one in the second channel. Users of the G_C with a probability of

$$p_{DTP} = \min\left\{1, \frac{2}{4}\right\} = 0.5$$

decide to send them data in this frame or not. Suppose that users U_1 and U_4 decided not to transmit, and users U_2 and U_5 decided to transmit and reselected channels 1 and 3, respectively. As a result, the “Success” event will occur in all three channels.

Consider the second frame. At the beginning of this frame, there are messages from users U_1 and U_4 from the previous frame, as well as new ones from users U_6, U_7 and U_8 . In the EP, users decide to transmit preambles or not with a probability equal

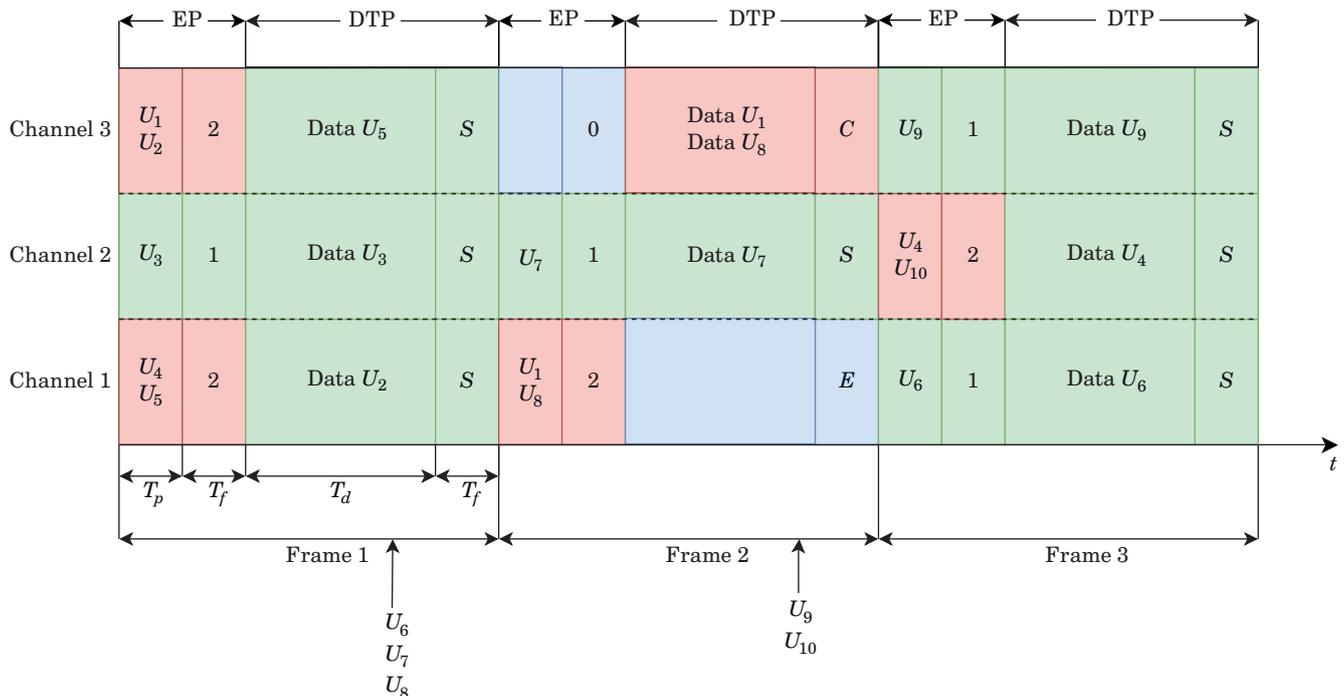


Fig. 3. Example of the lossless algorithm’s work

to $p_{EP} = \min\left\{1, \frac{3}{5}\right\} = 0.6$. Suppose that users U_4

and U_6 decided not to transmit, and users U_1, U_7 and U_8 decided to transmit. User U_7 chose the second channel, and users U_1 and U_8 chose the first channel. The BS, having estimated the number of preambles in each channel, will return the following information to all users $\{2, 1, 0\}$. Thus, user U_7 belongs to the G_S , and users U_1, U_8 to the G_C .

Then the DTP follows, in which the user U_7 transmits with probability one in the second channel. In this case, users of G_C transmit data with a

probability of $p_{DTP} = \min\left\{1, \frac{2}{2}\right\} = 1$. Let's assume

that users U_1 and U_8 have both re-selected the third channel. As a result, the "Empty" event will occur in the first channel, the "Success" event in the second channel, and the "Conflict" event in the third channel.

Consider the third frame. At the beginning of this frame there are messages from users U_1, U_4, U_6, U_8 from the previous frames, as well as new ones from users U_9 and U_{10} . Suppose that users U_4, U_6, U_9 and U_{10} decide to transmit in EP. User U_6 chose the first channel, users U_4 and U_{10} chose the second channel, and U_9 chose the third channel. BS, having estimated the number of preambles in each channel, will return the following information to all users $\{1, 2, 1\}$. Thus, users U_6 and U_9 belong to the G_S , and users U_4, U_{10} to the G_C .

This is followed by the DTP. Suppose that user U_{10} decided not to transmit, and user U_4 decided to transmit and re-selected the second channel. As a result, the "Success" event will occur in all three channels.

For a system with repeated transmissions, we introduce two characteristics [16]:

- the dependence of the average delay on input arrival rate $d(\lambda)$;
- the limiting input arrival rate at which the system works stably $\lambda_{cr} \triangleq \sup\{\lambda : d(\lambda) < \infty\}$.

Let us show how, for a system described by a set of assumptions, λ_{cr} can be calculated.

The number of active users during system operation can be described by the following recursive equation:

$$M_{t+1} = M_t - N_t + V_t,$$

where M_t and M_{t+1} are the number of users in the system at time t and $t + 1$, respectively; N_t is the number of users who logged out of the system at time t , $N_t \in \{0, 1, 2, \dots, K\}$, K is the number of channels in the system; V_t is the number of users who arrived at time t , distributed according to the Poisson law with the parameter λ .

It follows from this description that the sequence of random variables M_t defines a homogeneous irreducible aperiodic Markov chain.

To analyze stability, we calculate the mathematical expectation $E[N_t | M_t = m]$. It follows from the Foster criterion (see, for example, [17]) that λ_{cr} can be defined as the following limit:

$$\lambda_{cr} = \lim_{m \rightarrow \infty} \frac{E[N_t | M_t = m]}{K}. \quad (4)$$

Denote by $N_t^{(i)}$ the number of users leaving channel i at time t , then $N_t = N_t^{(1)} + N_t^{(2)} + \dots + N_t^{(K)}$. Then from (4) it follows:

$$\begin{aligned} \lambda_{cr} &= \lim_{m \rightarrow \infty} \frac{E[N_t^{(1)} + N_t^{(2)} + \dots + N_t^{(K)} | M_t = m]}{K} = \\ &= \lim_{m \rightarrow \infty} \frac{\sum_{i=1}^K E[N_t^{(i)} | M_t = m]}{K} = \\ &= \frac{\sum_{i=1}^K \lim_{m \rightarrow \infty} E[N_t^{(i)} | M_t = m]}{K}. \end{aligned}$$

It can be noted that for any value of m

$$\begin{aligned} E[N_t^{(1)} | M_t = m] &= \\ &= E[N_t^{(2)} | M_t = m] = \dots = E[N_t^{(K)} | M_t = m]. \end{aligned}$$

Then:

$$\lambda_{cr} = \lim_{m \rightarrow \infty} E[N_t^{(1)} | M_t = m].$$

Taking into account the fact that only message one user's can be successfully transmitted in one channel, then $E[N_t^{(1)} | M_t = m] = \Pr\{N_t^{(1)} = 1 | M_t = m\}$. Then:

$$\lambda_{cr} = \lim_{m \rightarrow \infty} \Pr\{N_t^{(1)} = 1 | M_t = m\}. \quad (5)$$

To calculate (5), we formulate and prove the following assertion.

Statement 1. Let there be m active users in some frame with number t . Then, as $m \rightarrow \infty$, each channel independently receives a Poisson input flow with intensity 1.

Proof: Let's consider the channel number i . Let $M_t = m$ and $m \geq K, m \geq G$, then all users decide to transmit with probability $p_{EP} = \frac{KG}{m}$ and those users who decide to transmit will choose the i -th chan-

nel with probability $p = \frac{1}{K}$. Denote by $Z_t^{(i)}$ the number of such users. The random variable $Z_t^{(i)}$ is distributed according to the Binomial distribution with parameters $\frac{G}{m}$ and m , then $E[Z_t^{(i)}] = \frac{G}{m}m = G$. If $m \rightarrow \infty$, then the Binomial distribution becomes a Poisson distribution with the same expectation.

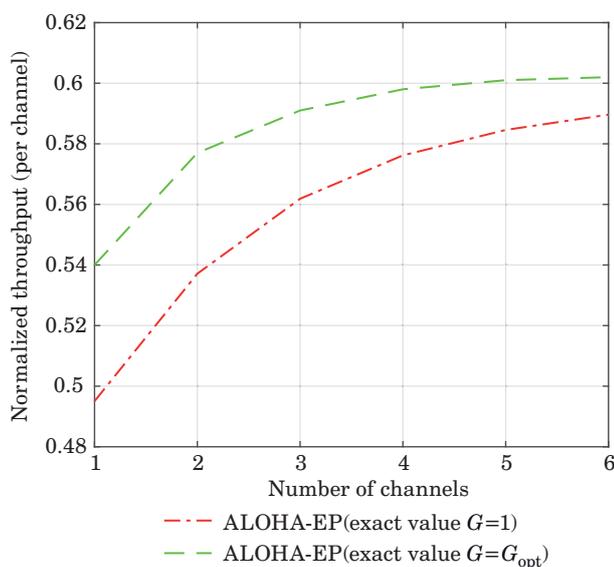
It follows from the statement that the value of the limit (5) can be calculated by substituting the value $\lambda = G$ into (3). Thus, using (3), we can calculate the value of λ_{cr} . The optimal value of G that maximizes the value of λ_{cr} is defined as

$$G_{opt} = \arg \max_{\lambda} T(\lambda).$$

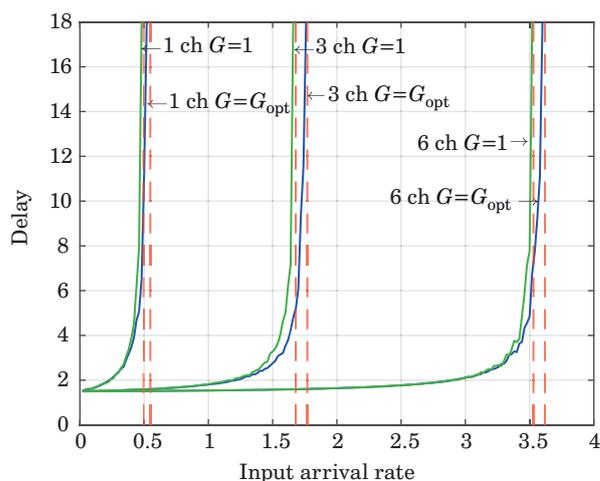
Figure 3 shows that the maximum limiting intensity with a small number of channels is achieved at the parameter $G_{opt} > 1$.

On Fig. 4 shows the results of calculating the critical input intensity up to which the system is stable according to formula (3) for the values of the number of channels 1, 2, 3, 4, 5 and 6. It is also shown that the optimal choice of the G parameter makes it possible to increase the throughput with a small number of channels. With an increase in the number of channels, this gain decreases, and with a large number of channels, the optimal value is $G = 1$.

To numerically calculate the average delay in the system, it is necessary to find the stationary distribution of the Markov chain with a countable number of states and cumbersome expressions for calculating the transition probabilities. Therefore, to illustrate the effect of the parameter G on the delay, a Java pro-



■ Fig. 4. Throughput on the number of channels



■ Fig. 5. Delay in system without losses at $G = 1$ and $G = opt$ from input arrival rate per channel

gram was developed and implemented that simulates the operation of the system, taking into account the assumptions described earlier. The results obtained by simulation modeling are presented in Fig. 5. The number of experiments was chosen such that the value of the confidence interval coincides with the thickness of the graph. As the throughput is approached, the delay increases and then tends to infinity, in which case the system is not stable.

Conclusion

The paper considers a model of a random multiple access system based on ALOHA with exploration phase from [12], where the calculation of the throughput per channel for an infinite number of channels was proposed. A method for analyzing the characteristics of the algorithm is proposed, which, in contrast to [12], makes it possible to calculate the throughput per channel for a given number of channels. It is shown that the maximum throughput with a finite number of channels is achieved at the input intensity $\lambda > 1$.

A modification of the algorithm for lossless operation is proposed by using repeated transmissions and changing the operation of the exploration phase to stabilize the operation of the algorithm. It is shown that the system provides stable operation when the input arrival rate does not exceed the maximum throughput. The dependence of delays in the system, obtained using simulation modeling, is also given.

The model considered in the paper and its modification is based on 3 unrealistic assumptions: an infinite number of preambles, the availability of information on the number of active users, and no time is spent on the exploration phase. These assumptions can be consistently discarded in further

studies. Can use the method of estimating the number of active users (see, for example, [18–20]). If the number of preambles is limited, then the results of the above analysis will be an upper bound on the throughput. Also in the future, it is necessary to take into account the duration of the exploration phase and the response of the base station.

References

1. Ratasuk R., Prasad A., Li Z., Ghosh A., & Uusitalo M. A. Recent advancements in M2M communications in 4G networks and evolution towards 5G. *18th Intern. Conf. on Intelligence in Next Generation Networks*, IEEE, 2015, pp. 52–57. doi:10.1109/ICIN.2015.7073806
2. Letaief K. B., Chen W., Shi Y., Zhang J., and Zhang Y. J. A. The roadmap to 6G: AI empowered wireless networks. *IEEE Communications Magazine*, 2019, vol. 57(8), pp. 84–90. doi:10.1109/MCOM.2019.1900271
3. Piran M. J., and Suh D. Y. Learning-driven wireless communications, towards 6G. *Intern. Conf. on Computing, Electronics & Communications Engineering (iCCECE)*, IEEE, 2019, pp. 219–224. doi:10.1109/iCCECE46942.2019.8941882
4. Wu Y., Gao X., Zhou S., Yang W., Polyanskiy Y., & Caire G. Massive access for future wireless communication systems. *IEEE Wireless Communications*, 2020, vol. 27(4), pp. 148–156. doi:10.1109/MWC.001.1900494
5. Shahab M. B., Abbas R., Shirvanimoghaddam M., and Johnson S. J. Grant-free non-orthogonal multiple access for iot: A survey. *IEEE Communications Surveys & Tutorials*, IEEE, 2020, vol. 22, no. 3, pp. 1805–1838. doi:10.1109/COMST.2020.2996032
6. Choi J. Noma-based random access with multichannel aloha. *IEEE Journal on Selected Areas in Communications*, IEEE, 2017, vol. 35, no. 12, pp. 2736–2743. doi:10.1109/JSAC.2017.2766778
7. Balevi E., Al Rabee F. T., and Gitlin R. D. ALOHA-noma for massive machine-to-machine IoT communication. *IEEE Intern. Conf. on Communications (ICC)*, IEEE, 2018, pp. 1–5. doi:10.1109/ICC.2018.8422892
8. Matveev N. V. The adaptive retransmission management in random multiple-access system with successive interference cancellation. *2018 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*, IEEE, 2018, pp. 1–5. doi:10.1109/WECONF.2018.8604323
9. Burkov A., Frolov A., and Turlikov A. Contention-based protocol with time division collision resolution. *2018 10th Intern. Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, IEEE, 2018, pp. 1–4. doi:10.1109/ICUMT.2018.8631217
10. Burkov A., Saveliev A., and Turlikov A. Upper and lower bound for non-blocking random multiple access algorithm with time division mode. *2019 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*, IEEE, 2019, pp. 1–7. doi:10.1109/WECONF.2019.8840590
11. Burkov A. A., Shneer S. V., Turlikov A. M. Lower bound for average delay in unblocked random access algorithm with orthogonal preambles. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 3, pp. 79–85. doi:10.31799/1684-8853-2020-3-79-85
12. Choi J. On improving throughput of multichannel ALOHA using preamble-based exploration. *Journal of Communications and Networks*, IEEE, 2020, vol. 22, no. 5, pp. 380–389. doi:10.1109/JCN.2020.000024
13. Bogatyrev V. A., Bogatyrev S. V., Bogatyrev A. V. Model and interaction efficiency of computer nodes based on transfer reservation at multipath routing. *2019 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*, IEEE, 2019, pp. 1–4. doi:10.1109/WECONF.2019.8840647
14. Bogatyrev V. A., Bogatyrev A. V., Bogatyrev S. V. Redundant servicing of a flow of heterogeneous requests critical to the total waiting time during the multi-path passage of a sequence of info-communication nodes. *Intern. Conf. on Distributed Computer and Communication Networks*, Springer, Cham, 2020, pp. 100–112. doi:10.1007/978-3-030-66471-8_9
15. Bogatyrev V. A., Bogatyrev A. V., Bogatyrev S. V. Redundant multi-path service of a flow heterogeneous in delay criticality with defined node passage paths. *Journal of Physics: Conference Series*, IOP Publishing, 2021, vol. 1864, no. 1, p. 012094. doi:10.1088/1742-6596/1864/1/012094
16. Tsybakov B. Survey of USSR contributions to random multiple-access communications. *IEEE Transactions on Information Theory*, 1985, vol. 31, no. 2, pp. 143–165. doi:10.1109/TIT.1985.1057023
17. Tsybakov B. S., and Mikhailov V. A. Free synchronous packet access in a broadcast channel with feedback. *Problemy peredachi informatsii* [Problems Information Transmission], 1978, vol. 14, no. 4, pp. 259–280 (In Russian).
18. Galinina O., Turlikov A., Andreev S., and Koucheryavy Y. Stabilizing multichannel slotted aloha for machine-type communications. *2013 IEEE Intern. Symp. on Information Theory*, IEEE, 2013, pp. 2119–2123. doi:10.1109/ISIT.2013.6620600
19. Yu J., Zhang P., Chen, L., Liu J., Zhang R., Wang K., & An J. Stabilizing frame slotted ALOHA-based IoT

Financial support

The paper was prepared with the financial support of the Russian Science Foundation, project No. 22-19-00305 “Spatial-temporal stochastic models of wireless networks with a large number of users”.

systems: A geometric ergodicity perspective. *IEEE Journal on Selected Areas in Communications*, IEEE, 2020, vol. 39(3), pp. 714–725. doi:10.1109/JSAC.2020.3018795

20. Liu J., Seo J. B., Jin H. Online transmission control for random access with multipacket reception and reservation. *IEEE Internet of Things Journal*, IEEE, 2022. doi:10.1109/JIOT.2022.3188280

УДК 004.728.3.057.4

doi:10.31799/1684-8853-2022-5-49-59

EDN: KGVYDI

Анализ и стабилизация многоканального алгоритма ALOHA с использованием фазы исследования на основе преамбул

А. А. Бурков^а, ассистент, orcid.org/0000-0002-0920-585X

Р. О. Рачугин^а, магистрант, orcid.org/0000-0001-5813-3867

А. М. Тюрликов^а, доктор техн. наук, профессор, orcid.org/0000-0001-7132-094X, turlikov@vu.spb.ru

^аСанкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ

Введение: в настоящее время активно используются устройства интернета вещей, функционирующие в рамках сценариев массовых машинных коммуникаций. Взаимодействие устройств осуществляется алгоритмами случайного множественного доступа, имеющими ограниченную пропускную способность. Для увеличения пропускной способности можно использовать ортогональные преамбулы в классе алгоритмов ALOHA. **Цель:** провести анализ алгоритмов на базе ALOHA, использующих фазу исследования, и вычислить характеристики для алгоритма с потерями и без потерь при конечном числе каналов. **Результаты:** описана модель системы случайного доступа для передачи данных по общему каналу связи с использованием ортогональных преамбул и фазы исследования. Получена формула для численного расчета пропускной способности на канал алгоритма с потерями при бесконечном числе преамбул и заданном конечном числе каналов. Представлены результаты расчета для нескольких значений числа независимых каналов. Предложена и описана модификация алгоритма, использующая фазу исследования и повторные передачи. Данная система может работать без потерь. Для этой системы приведен анализ предельной пропускной способности, до которой система работает стабильно. Также показаны значения средней задержки для алгоритма, полученные имитационным моделированием. При уменьшении числа доступных преамбул данные результаты можно использовать в качестве верхней границы пропускной способности системы. **Практическая значимость:** полученные результаты позволяют оценить потенциальные возможности увеличения пропускной способности систем случайного множественного доступа в сетях 6G за счет применения фазы исследования.

Ключевые слова – многоканальная ALOHA, предельная интенсивность выходного потока, случайный множественный доступ, стабильность, массовая межмашинная связь, интернет вещей.

Для цитирования: Burkov A. A., Rachugin R. O., Turlikov A. M. Analyzing and stabilizing multichannel ALOHA with the use of the preamble-based exploration phase. *Информационно-управляющие системы*, 2022, № 5, с. 49–59. doi:10.31799/1684-8853-2022-5-49-59, EDN: KGVYDI

For citation: Burkov A. A., Rachugin R. O., Turlikov A. M. Analyzing and stabilizing multichannel ALOHA with the use of the preamble-based exploration phase. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2022, no. 5, pp. 49–59. doi:10.31799/1684-8853-2022-5-49-59, EDN: KGVYDI

УВАЖАЕМЫЕ АВТОРЫ!

Научные базы данных, включая Scopus и Web of Science, обрабатывают данные автоматически. С одной стороны, это ускоряет процесс обработки данных, с другой – различия в транслитерации ФИО, неточные данные о месте работы, области научного знания и т. д. приводят к тому, что в базах оказывается несколько авторских страниц для одного и того же человека. В результате для всех по отдельности считаются индексы цитирования, что снижает рейтинг ученого.

Для идентификации авторов в сетях Thomson Reuters проводит регистрацию с присвоением уникального индекса (ID) для каждого из авторов научных публикаций.

Процедура получения ID бесплатна и очень проста, есть возможность провести регистрацию на 12 языках, включая русский (чтобы выбрать язык, кликните на зеленое поле сверху справа на стартовой странице): <https://orcid.org>

АДЁРКИНА
Анастасия
Александровна



Аспирант, преподаватель кафедры радиотехники радиофизического факультета Национального исследовательского Нижегородского государственного университета им. Н. И. Лобачевского.

В 2018 году с отличием окончила магистратуру Национального исследовательского Нижегородского государственного университета им. Н. И. Лобачевского по специальности «Информационные процессы и системы».

Является автором восьми научных публикаций.

Область научных интересов — цифровая обработка сигналов, теория распространения радиосигналов, применение методов машинного обучения в радиофизических задачах.

Эл. адрес: aaderkina@mail.ru

БЕЗЗАТЕЕВ
Сергей
Валентинович



Заведующий кафедрой технологий защиты информации и технологической безопасности Санкт-Петербургского государственного университета аэрокосмического приборостроения.

В 1980 году окончил Ленинградский институт авиационного приборостроения по специальности «Автоматизированные системы управления».

В 2011 году защитил диссертацию на соискание ученой степени доктора технических наук. Является автором более 40 научных публикаций.

Область научных интересов — теория информации, теория кодирования, системы информационной безопасности.

Эл. адрес: bsv@aanet.ru

БУРКОВ
Артем
Андреевич



Ассистент кафедры инфокоммуникационных технологий и систем связи Санкт-Петербургского государственного университета аэрокосмического приборостроения.

В 2017 году окончил с отличием Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Инфокоммуникационные технологии и системы связи».

Является автором 23 научных публикаций.

Область научных интересов — беспроводные системы передачи данных, алгоритмы случайного множественного доступа, системы с гибридной решаемой обратной связью, цифровая обработка изображений, сжатие данных.

Эл. адрес: a.burkov@k36.org

ДВОЙНИКОВА
Анастасия
Александровна



Младший научный сотрудник лаборатории речевых и мультимедийных интерфейсов Санкт-Петербургского Федерального исследовательского центра РАН. В 2021 году окончила Университет ИТМО по специальности «Речевые информационные системы».

Является автором десяти научных публикаций.

Область научных интересов — аффективные вычисления, анализ вовлеченности, анализ эмоций, сентимент-анализ, компьютерная паралингвистика.

Эл. адрес: dvoynikova.a@ias.spb.ru

ИСАЕВ
Сергей
Владиславович



Доцент, заместитель директора по научной работе Института вычислительного моделирования СО РАН — обособленного подразделения Федерального исследовательского центра «Красноярский научный центр СО РАН», доцент Сибирского федерального университета.

В 1993 году окончил Красноярский государственный университет по специальности «Математика».

В 1999 году защитил диссертацию на соискание ученой степени кандидата технических наук. Является автором более 100 научных публикаций.

Область научных интересов — информационная безопасность, распределенные информационные системы, интеллектуальные системы.

Эл. адрес: si@krasn.ru

ИСАЕВА
Ольга
Сергеевна



Старший научный сотрудник отдела прикладной информатики Института вычислительного моделирования СО РАН — обособленного подразделения Федерального исследовательского центра «Красноярский научный центр Сибирского отделения Российской академии наук».

В 1998 году окончила Красноярский государственный университет по специальности «Прикладная математика».

В 2004 году защитила диссертацию на соискание ученой степени кандидата технических наук. Является автором 95 научных публикаций и 13 свидетельств о регистрации программ для ЭВМ.

Область научных интересов — методы искусственного интеллекта, анализ данных, цифровые двойники.

Эл. адрес: isaeva@icm.krasn.ru

**КАГИРОВ
Ильдар
Амирович**



Научный сотрудник лаборатории речевых и многомодальных интерфейсов Санкт-Петербургского Федерального исследовательского центра РАН. В 2008 году окончил филологический факультет Санкт-Петербургского государственного университета по специальности «Лингвистика». Является автором более 30 научных публикаций и одного патента на изобретения. Область научных интересов — лингвистика жестовых языков, человеко-машинное взаимодействие, синтаксис и грамматическая семантика естественных языков, корпусная лингвистика. Эл. адрес: kagirov@iias.spb.su

**КАРПОВ
Алексей
Анатольевич**



Главный научный сотрудник, руководитель лаборатории речевых и многомодальных интерфейсов Санкт-Петербургского Федерального исследовательского центра РАН. В 2002 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Вычислительные машины, комплексы, системы и сети». В 2013 году защитил диссертацию на соискание ученой степени доктора технических наук. Является автором более 350 научных публикаций, включая четыре монографии и пять патентов. Область научных интересов — многомодальные интерфейсы и системы, речевые технологии, автоматическое распознавание и синтез речи, компьютерная паралингвистика. Эл. адрес: karpov@iias.spb.su

**КУЛЯСОВ
Никита
Владимирович**



Программист первой категории отдела информационно-телекоммуникационных технологий Института вычислительного моделирования СО РАН — обособленного подразделения Федерального исследовательского центра «Красноярский научный центр Сибирского отделения Российской академии наук». В 2021 году окончил Сибирский федеральный университет по специальности «Информатика и вычислительная техника». Является автором четырех научных публикаций и двух свидетельств о регистрации программы для ЭВМ. Область научных интересов — высоконагруженные системы и приложения. Эл. адрес: razor@icm.krasn.ru

**ПИМЕНОВ
Андрей
Владимирович**



Студент факультета безопасности информационных технологий Университета ИТМО, Санкт-Петербург. Является автором двух научных публикаций. Область научных интересов — обеспечение информационной безопасности в информационных системах, применение технологий блокчейн и машинного обучения для обеспечения информационной безопасности. Эл. адрес: tik11994@mail.ru

**РАЧУГИН
Роман
Олегович**



Магистрант кафедры инфокоммуникационных систем Санкт-Петербургского государственного университета аэрокосмического приборостроения. В 2021 году окончил бакалавриат Санкт-Петербургского государственного университета аэрокосмического приборостроения по специальности «Инфокоммуникационные технологии и системы связи». Область научных интересов — теория разрядных вычислений, методы проектирования спецпроцессоров для систем контроля и управления, опто-информационные системы. Эл. адрес: rro1699@gmail.com

**СИНИЦЫН
Александр
Сергеевич**



Студент радиофизического факультета Национального исследовательского Нижегородского государственного университета им. Н. И. Лобачевского. Является автором двух научных публикаций. Область научных интересов — цифровая обработка сигналов, теория распространения радиосигналов, применение методов машинного обучения в радиофизических задачах. Эл. адрес: sinitsynalexandr@mail.ru

**ТУМКОВСКИЙ
Сергей
Ростиславович**



Профессор, заместитель директора Московского института электроники и математики им. А. Н. Тихонова Национального исследовательского университета «Высшая школа экономики», лауреат премии Правительства РФ в области науки и техники. В 1981 году окончил Московский институт электронного машиностроения по специальности «Автоматика и телемеханика». В 2007 году защитил диссертацию на соискание ученой степени доктора технических наук. Является автором более 140 научных публикаций и 11 свидетельств о регистрации ПО. Область научных интересов – системы автоматизации проектирования, методы математического моделирования физических процессов в радиоэлектронных средствах.
Эл. адрес: STumkovskiy@hse.ru

**ТЮРЛИКОВ
Андрей
Михайлович**



Профессор, заведующий кафедрой инфокоммуникационных технологий и систем связи Санкт-Петербургского государственного университета аэрокосмического приборостроения. В 1980 году окончил Ленинградский институт авиационного приборостроения по специальности «Информационные системы управления». В 2011 году защитил диссертацию на соискание ученой степени доктора технических наук. Является автором более 150 научных публикаций. Область научных интересов – многоабонентные системы связи, системы дистанционного обучения, протоколы передачи данных в реальном масштабе времени, алгоритмы сжатия видеoinформации.
Эл. адрес: turlikov@k36.org

**ТЫРЬШКИНА
Евгения
Сергеевна**



Старший разработчик хранилища данных группы компаний «VK», аспирант Московского института электроники и математики им. А. Н. Тихонова Национального исследовательского университета «Высшая школа экономики». В 2018 году окончила магистратуру Национального исследовательского университета «Высшая школа экономики» по специальности «Информатика и вычислительная техника». Является автором девяти научных публикаций. Область научных интересов – обработка больших данных, системы управления данными, построение хранилищ данных.
Эл. адрес: tyryshkina_evgeniya@mail.ru

**ФЕДОРОВ
Иван
Романович**



Аспирант факультета информационной безопасности Университета ИТМО, Санкт-Петербург. В 2018 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Информатика и вычислительная техника». Является автором 12 научных публикаций и одного патента на изобретения. Область научных интересов – информационная безопасность, сети и телекоммуникации, блокчейн.
Эл. адрес: ivanfedorov@itmo.ru